

Method and system for pattern recognition based on dynamically constructing a subset of reference vectors

Publication number: JP10508391T

Publication date: 1998-08-18

Inventor:

Applicant:

Classification:

- international: **G10L15/08; G10L15/14; G10L15/00; (IPC1-7): G10L3/00**

- european: G10L15/08

Application number: JP19960510050T 19960826

Priority number(s): EP19950202317 19950828; WO1996IB00849 19960826

Also published as:

WO9708685 (A3)
WO9708685 (A2)
EP0788648 (A3)
EP0788648 (A2)
US5933806 (A1)

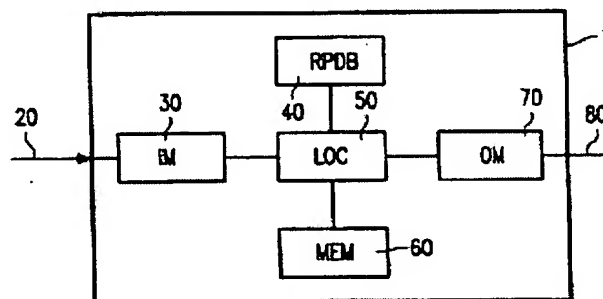
more >>

Report a data error he

Abstract not available for JP10508391T

Abstract of corresponding document: **US5933806**

A system and method are used for recognising a time-sequential input pattern (20), which is derived from a continual physical quantity, such as speech. The system has input means (30), which accesses the physical quantity and therefrom generates a plurality of input observation vectors. The input observation vectors represent the input pattern. A reference pattern database (40) is used for storing a plurality of reference patterns. Each reference pattern includes a sequence of reference units, where each reference unit is represented by at least one associated reference vector $+E, ovs \mu +EE a$ in a set $\{+E, ovs \mu +EE a\}$ of reference vectors. A localizer (50) is used for locating among the reference patterns stored in the reference pattern database (40), a recognised reference pattern, which corresponds to the input pattern. The locating includes selecting a subset $\{+E, ovs \mu +EE s\}$ of reference vectors from said set $\{+E, ovs \mu +EE a\}$ for each input observation vector $+E, ovs o+EE$ by calculating a measure of dissimilarity between the input observation vector $+E, ovs o+EE$ and each reference vector of the set $\{+E, ovs \mu +EE a\}$. A number of reference vectors $+E, ovs \mu +EE a$, whose measures of dissimilarity with said input observation vector $+E, ovs o+EE$ are the smallest, are used as the subset $\{+E, ovs \mu +EE s\}$ of reference vectors. The reference vectors of the subset $\{+E, ovs \mu +EE s\}$ are used to calculate for each reference pattern, a pattern similarity score. The recognised pattern is one of the reference patterns for which an optimum of the pattern similarity scores is calculated. Output means (70) are used for outputting the recognised pattern.



Data supplied from the esp@cenet database - Worldwide

(19) 日本国特許庁 (J P)

(12) 公表特許公報 (A)

(11) 特許出願公表番号

特表平10-508391

(43) 公表日 平成10年(1998) 8月18日

(51) Int.Cl.⁶

G10L 3/00

識別記号

531

F I

G10L 3/00

531 E

審査請求 未請求 予備審査請求 未請求(全 46 頁)

(21) 出願番号 特願平9-510050
 (86) (22) 出願日 平成8年(1996) 8月28日
 (85) 翻訳文提出日 平成9年(1997) 4月25日
 (86) 国際出願番号 PCT/JP96/00849
 (87) 国際公開番号 WO97/08685
 (87) 国際公開日 平成9年(1997) 3月6日
 (31) 優先権主張番号 95202317.4
 (32) 優先日 1995年8月23日
 (33) 優先権主張国 オランダ (NL)
 (81) 指定国 EP(AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), JP

(71) 出願人 フィリップス エレクトロニクス ネムローゼ フェンノートシャップ
 オランダ国 5621 ペーアー アインドーフェン フルーネヴァウツウェッハ 1
 (72) 発明者 ベイヤライン ベーター
 ドイツ連邦共和国 52078 アーヘン エレンドルフアーシュトラッセ 87
 (72) 発明者 ウルリッヒ マインハルト ディーター
 ドイツ連邦共和国 50935 ケルン カイザー セシャーシュトラッセ 1
 (74) 代理人 弁理士 杉村 暁秀 (外6名)

(54) 【発明の名称】 基準ベクトルの部分集合の動的な形成に基づくパターン認識の方法及びシステム

(57) 【要約】

このシステム及び方法は、音声のように連続的な物理量から導出される時系列入力パターン(20)を認識するために用いられる。このシステムは、物理量にアクセスしそれから複数の入力観察ベクトルを発生する認識手段(30)を具える。入力観察ベクトルが入力パターンを表現する。基準パターンデータベース(40)が、複数の基準パターンを格納するために用いられる。各基準パターンは基準ユニットの列からなり、各基準ユニットは、基準ベクトルの集合{ μ_i }中の少なくとも1つの対応する基準ベクトル μ_{ij} によって表現される。局所化装置(50)が、基準パターンデータベース(40)中に格納された基準パターンの間に、入力パターンに対応し認識された基準パターンを局所化するために用いられる。

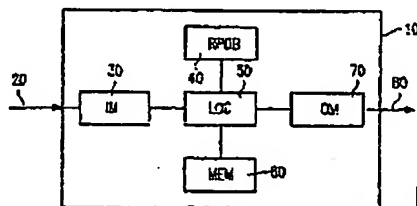


FIG. 3

(2)

特表平10-508391

【特許請求の範囲】

1. 入力パターンを表す物理量及びそれから発生する複数の入力観察ベクトルにアクセスし、

入力パターンに対応して認識された基準パターンを複数の基準パターンの間に局所化を行い、少なくとも1つの基準パターンが基準ユニットの列であり、

各基準ユニットが基準ベクトルの集合 $\{\overline{\mu}_i\}$ の少なくとも1つの対応する基準ベクトル $\overline{\mu}_i$ によって表現され、前記局所化は、各入力観察ベクトル \overline{o} について前記集合 $\{\overline{\mu}_i\}$ から基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ を選択すること、及び、前記入力観察ベクトル \overline{o} と前記部分集合 $\{\overline{\mu}_j\}$ の各基準ベクトル $\overline{\mu}_j$ 、

との間のベクトル類似度を計算することによって行う、連続的な物理量から導出される入力パターンを認識する方法において、

各入力観察ベクトル \overline{o} についての基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ の選択が、前記入力観察ベクトル \overline{o} と前記集合 $\{\overline{\mu}_j\}$ の各基準ベクトルとの間の非類似度の大きさの計算、及び、基準ベクトルの前記部分集合 $\{\overline{\mu}_j\}$ として前記入力観察ベクトル \overline{o} との非類似度の大きさが最小である多数の基準ベクトル $\overline{\mu}_j$ 、

の使用を含むことを特徴とする方法。

2. 前記方法が、

各基準ベクトル $\overline{\mu}_i$ を量子化して量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ とし、且つ基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ の選択が、各入力観察ベクトル \overline{o} について

次のステップ、即ち

入力観察ベクトル \overline{o} を量子化して量子化入力観察ベクトル $\overline{R}(\overline{o})$ とするス

テップ、

各量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ に対する前記量子化入力観察ベクトル $\overline{R}(\overline{o})$ の距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算するステップ、及び

前記入力観察ベクトル \overline{o} と前記基準ベクトル $\overline{\mu}_j$ との間の前記非類似度の大きさとして、前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_j))$ を使用するステップ

を含むことを特徴とする請求項1に記載の方法。

3. ベクトル \overline{x} (基準ベクトル $\overline{\mu}_i$ 、又は観察ベクトル \overline{o}) を量子化ベクトル $\overline{R}(\overline{x})$

(3)

特表平10-508391

\overline{x})とする量子化が、符号ベクトルの各成分を、ベクトル \overline{x} の対応する成分が負値を持つならば第1二値 b_1 に割当て、ベクトル \overline{x} の対応する成分が正値を

持つならば第2二値 b_2 に割当てる方法で、二値への割当てによる符号ベクトル $\overline{S}(\overline{x})$ の計算を含むことを特徴とする請求項2に記載の方法。

4. 前記量子化が、更に、ベクトル \overline{x} のL₁ノルムの計算、及び、前記ノルムと前記符号ベクトル $\overline{S}(\overline{x})$ との乗算を含むことを特徴とする請求項3に記載の

方法。

5. 前記量子化が、更に、符号ベクトル $\overline{S}(\overline{x})$ のベクトル \overline{x} のディメンション

による除算と $1/r$ 乗を含むことを特徴とする請求項4に記載の方法。

6. 前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_s))$ の計算が、ベクトル $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_s)$ とのハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_s))$ の計算を含むことを特徴とする請求項3に記載の方法。

7. 前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_s))$ の計算が、

各ベクトル $\overline{\mu}_s$ のL₁ノルム $\|\overline{\mu}_s\|_1$ の計算、及び

各入力観察ベクトル \overline{o} について、

ベクトル \overline{o} のL₁ノルム $\|\overline{o}\|_1$ の計算、及び

ベクトル $\overline{S}(\overline{o})$ の各ベクトル $\overline{S}(\overline{\mu}_s)$ に対するハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_s))$ の計算

を含むことを特徴とする請求項4又は5に記載の方法。

8. ベクトル $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_s)$ との前記ハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_s))$ の計算が、

$\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_s)$ との対応する成分の二値XOR値を、差分ベ

クトルの各成分に割当てることにより差分ベクトルを計算すること、

前記差分ベクトルで値1を持つ成分の数を計算することにより差分数を求めること、及び

(4)

特表平10-508391

ハミング距離として前記差分数を用いること

を含むことを特徴とする請求項6又は7に記載の方法。

9. 前記方法が、

0又は1の二値を持つ成分からなる各N次元ベクトルについて、値1を持つ成分の数を表す対応する数を特定するテーブルを作ること、及び

前記差分ベクトルを前記テーブルに位置させ且つ前記対応する数をハミング距離として用いて前記差分数を求めること

を含むことを特徴とする請求項8に記載の方法。

10. 前記方法が、予め定められた入力観察ベクトル \bar{o} についての基準ベクトルの

部分集合を選択し、続く多数の観察ベクトルに対して同一の部分集合を用いるようにすることを特徴とする請求項1乃至9のいずれか1項に記載の方法。

11. 1つの入力観察ベクトル \bar{o} について基準ベクトルの前記部分集合を選択し、

表現されない各基準ユニットについて、代表基準ベクトルを前記部分集合に加算することにより、各基準ユニットが前記部分集合の少なくとも1つの基準ベクトルで表現されるようにすることを特徴とする請求項1乃至10のいずれか1項に記載の方法。

12. 各基準ユニットが前記部分集合の少なくとも1つの基準ベクトルで表現されるようにし、代表基準ベクトルとして前記基準ユニットを表現し且つ前記入力観察ベクトル \bar{o} に対して最小距離を持つ部分集合から基準ベクトルを選択する

ことにより、各基準ユニットについて前記代表基準ベクトルを選択することを特徴とする請求項11に記載の方法。

13. 入力パターンを表す物理量及びそれから発生する複数の入力観察ベクトルにアクセスする入力手段、

少なくとも1つの基準パターンが一連の基準ユニットからなり、各基準ユニ

(5)

特表平10-508391

ットが少なくとも1つの対応する基準ベクトルの集合 $\{\overline{\mu}_i\}$ の基準ベクトル $\overline{\mu}_i$ によった表現されるように構成され、この基準パターンを複数格納する基

準パターンデータベース、

前記入力パターンに対応する認識した基準パターンを、前記基準パターンデータベースに格納された複数の基準パターンの間に位置決めする局所化が、各入力観察ベクトル \overline{o} について前記集合 $\{\overline{\mu}_i\}$ から基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ を選択すること、及び前記入力観察ベクトル \overline{o} と前記部分集合 $\{\overline{\mu}_j\}$ の各基準ベクトル $\overline{\mu}_j$ との間のベクトル類似度を計算することを含み、そのた

めの局所化装置、及び

前記認識パターンを出力する出力手段

を具える連続的な物理量から導出される時間系列の入力パターンを認識するシステムにおいて、

各入力観察ベクトル \overline{o} についての基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ の選択が、前記入力観察ベクトル \overline{o} と前記集合 $\{\overline{\mu}_i\}$ の各基準ベクトルとの間の非類似度の大きさの計算、及び、基準ベクトルの前記部分集合 $\{\overline{\mu}_j\}$ として前記入力観察ベクトル \overline{o} との非類似度の大きさが最小である多数の基準ベクトル $\overline{\mu}_j$ 、

の使用を含むことを特徴とするシステム。

14. 前記基準パターンデータベースが、更に、各基準ベクトル $\overline{\mu}_i$ について量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ を格納し、及び
- 基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ の選択が、各入力観察ベクトル \overline{o} について

次のステップ、即ち

入力観察ベクトル \overline{o} を量子化して量子化入力観察ベクトル $\overline{R}(\overline{o})$ とするス

テップ、

(5)

特表平10-508391

各量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ に対する前記量子化入力観察ベクトル $\overline{R}(\overline{o})$ の距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算するステップ、及び

前記入力観察ベクトル \overline{o} と前記基準ベクトル $\overline{\mu}_i$ との間の前記非類似度の大きさとして、前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を使用するステップ

を含むことを特徴とする請求項13に記載のシステム。

15. ベクトル \overline{x} (基準ベクトル $\overline{\mu}_i$ 、又は観察ベクトル \overline{o}) について量子化ベクトル $\overline{R}(\overline{x})$ が符号ベクトル $\overline{S}(\overline{x})$ に比例し、前記符号ベクトル $\overline{S}(\overline{x})$ の各成分が、ベクトル \overline{x} の対応する成分が負値を持つならば第1二値 $b1$ であり、ベクトル \overline{x} の対応する成分が正値を持つならば第2二値 $b2$ であることを特徴

とする請求項14に記載のシステム。

16. 前記量子化ベクトル $\overline{R}(\overline{x})$ がベクトル \overline{x} の L_1 ノルムに比例することを特

徴とする請求項15に記載のシステム。

17. 前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ の計算が、ベクトル $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_i)$ とのハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ の計算を含むことを特徴とする請求項15に記載のシステム。

18. 各基準ベクトル $\overline{\mu}_i$ について基準パターンデータベースが基準ベクトル $\overline{\mu}_i$ の L_1 ノルム $\|\overline{\mu}_i\|_1$ を含み、

前記局所化装置が、各入力観察ベクトル \overline{o} について、

ベクトル \overline{o} の L_1 ノルム $\|\overline{o}\|_1$ 、及び、ベクトル $\overline{S}(\overline{o})$ の各ベクトル $\overline{S}(\overline{\mu}_i)$ に対するハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ を計算し、及び

L_1 ノルム $\|\overline{o}\|_1$ 及びハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ を、前記基準パターンデータベースに格納された L_1 ノルム $\|\overline{\mu}_i\|_1$ と組合せて、距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算する

ことを特徴とする請求項16に記載のシステム。

19. ベクトル $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_i)$ との前記ハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ の計算が、 $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_i)$ との対応する成分の二値XOR値を、差分ベ

(

(7)

特表平10-508391

クトルの各成分に割当てることにより差分ベクトルを計算すること、

前記差分ベクトルで値1を持つ成分の数を計算することにより差分数を求めること、及び

ハミング距離として前記差分数を用いること

を含むことを特徴とする請求項17又は18に記載のシステム。

20. 0又は1の二値を持つ成分からなる各N次元ベクトルについて、値1を持つ成分の数を表す対応する数を特定するテーブルを格納するメモリーを具え、及び

前記差分ベクトルを前記テーブルに位置させ且つ前記対応する数をハミング距離として用いて前記差分数を求めること

を含むことを特徴とする請求項19に記載のシステム。

21. 前記局所化装置が、予め定められた入力観察ベクトル \vec{o} についての基準ベク

トルの第1対応部分集合を選択し、続く多数の観察ベクトルに対して基準ベクトルの前記対応部分集合として前記第1部分集合を用いることを特徴とする請求項13乃至20のいずれか1項に記載のシステム。

22. 前記局所化装置が、1つの入力観察ベクトル \vec{o} について基準ベクトルの前記

部分集合を選択し、表現されない各基準ユニットについて、代表基準ベクトルを前記部分集合に加算することにより、各基準ユニットが前記部分集合の少なくとも1つの基準ベクトルで表現されるようにすることを特徴とする請求項13乃至21のいずれか1項に記載のシステム。

23. 前記局所化装置が、各基準ユニットが前記部分集合の少なくとも1つの基準ベクトルで表現されるようにし、代表基準ベクトルとして前記基準ユニットを表現し且つ前記入力観察ベクトル \vec{o} に対して最小距離を持つ部分集合から基準

ベクトルを選択することにより、各基準ユニットについて前記代表基準ベクトルを選択することを特徴とする請求項22に記載のシステム。

(8)

特表平10-508391

【発明の詳細な説明】

基準ベクトルの部分集合の動的な形成に基づくパターン認

識の方法及びシステム

本発明は、入力パターンを表す物理量及びそれから発生する複数の入力観察ベクトルにアクセスし、

入力パターンに対応して認識された基準パターンを複数の基準パターンの間に局所化を行い、少なくとも1つの基準パターンが基準ユニットの列であり、各基準ユニットが基準ベクトルの集合 $\{\overline{\mu}_i\}$ の少なくとも1つの対応する基準ベクトル $\overline{\mu}_i$ によって表現され、前記局所化は、各入力観察ベクトル \overline{o} について前記集合 $\{\overline{\mu}_i\}$ から基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ を選択すること、及び、前記入力観察ベクトル \overline{o} と前記部分集合 $\{\overline{\mu}_j\}$ の各基準ベクトル $\overline{\mu}_j$ との間のベクトル類似度を計算することによって行う、連続的な物理量から導出される入力パターンを認識する方法に関する。

本発明は、更に、連続的な物理量から導出される時間系列の入力パターンを認識するシステムであって、

入力パターンを表す物理量及びそれから発生する複数の入力観察ベクトルにアクセスする入力手段、

少なくとも1つの基準パターンが一連の基準ユニットからなり、各基準ユニットが少なくとも1つの対応する基準ベクトルの集合 $\{\overline{\mu}_i\}$ の基準ベクトル $\overline{\mu}_i$ 。

によった表現されるように構成され、この基準パターンを複数格納する基準パターンデータベース、

前記入力パターンに対応する認識した基準パターンを、前記基準パターンデータベースに格納された複数の基準パターンの間に位置決めする局所化が、各入力観察ベクトル \overline{o} について前記集合 $\{\overline{\mu}_i\}$ から基準ベクトルの部分集合 $\{\overline{\mu}_j\}$ を選択すること及び前記入力観察ベクトル \overline{o} と前記部分集合 $\{\overline{\mu}_j\}$ の各基準ベクトル $\overline{\mu}_j$ との間のベクトル類似度を計算することを含み、そのための局所化

置、及び

(9)

特表平10-508391

前記認識パターンを出力する出力手段

を具えるシステムに関する。

例えば音声又は画像のような、連続的な物理量から導出される時間系列の入力パターンの認識は重要性を増している。特に、音声認識は、例えば電話及び通信（種々の自動化サービス）、オフィス及びビジネスシステム（データ入力）、製造（製造プロセスのハンドフリーモニター）、医療（レポートの注解）、ゲーム（音声入力）、自動車の機能の音声制御及び身体障害者による音声制御の領域で最近広く利用されるようになった。連続的な音声認識について、図1に示すように、以下のような信号処理ステップが通常用いられる。[L.Rabiner"A Tutorial on Hidden Markov Model and Selected Application in Speech Recognition", Proceeding of the IEEE, 第77巻第2号, 1989年2月参照]

・特徴分析：音声入力信号がスペクトルの及び／又は時間的に分析され、特徴の代表ベクトル（観察ベクトル \vec{o} ）を計算する。典型的には、音声信号がディ

ジタル化（例えば6.67kHzのレートでサンプリングされる）され、例えば前強調を用いて前処理される。音声信号の連続サンプルが、例えば32msecに対応するフレームにグループ化（ブロック化）される。連続するフレームは部分的に例えば16msec間重なる。各フレームについて、特徴の代表ベクトル（観察ベクトル \vec{o} ）

を計算するために、線形予測符号化（LPC）スペクトル分析法がよく用いられる。この特徴ベクトルは例えば24、32又は63成分（特徴空間ディメンション）を持つ。

・ユニットマッチングシステム：観察ベクトルを音声認識ユニットの内容にマッチングさせる。音声認識のユニットとして種々のフォームが用いられる。幾つかのシステムが、例えば単音、二連音又は音節のような言語学的なサブワードユニット及び例えばフェネン及びフェノンのような二次ユニットを用いている。他のシステムはユニットとして全ワード又はワードのグループを用いる。いわゆる隠蔽マルコフモデル（HMM）が統計モデル音声信号に対して広く用いられている。このモデルを用いると、各ユニットがHMMによって典型的に特徴付けられ

(10)

特表平10-508391

る。この場合のパラメータは学習のための音声データの集合から見積もられる。
例えば10,000乃至60,000ワードを含む大規模語彙音声認識システムにおいては、

より多くのユニットについて適切にHMMを学習するためには多数の学習データが必要になるため、通常、例えば40個のサブワードユニットの限定された集合が用いられる。ユニットマッチングシステムは、観察ベクトルを全ての音声認識ユニットのシーケンスにマッチングさせ、ベクトルとシーケンスとの間のマッチングの可能性を与える。マッチングに例えば次のような制約をおくことができる。

－辞書による復号化：サブワードを用いる場合、発音辞書は、ワードがどのようにしてサブワードから作られるかを記述する。ユニットマッチングシステムによって探索されたサブワードユニットの可能なシーケンスが、次に辞書中のシーケンスに作り上げられる。

－文章論的分析：他の制約がユニットマッチングシステムにおかれ、それにより探索されたパスは、ワードを構成（辞書による復号化）し、そのためにワードがワード文法によって特定される適切なシーケンスにある音声ユニットに対応するパスである。

不連続マルコフ過程は、いつでもN個の個別状態の集合の1つであるシステムを記述する。正規の時刻にシステムはその状態に対応する確率の集合に従って状態を変える。不連続マルコフ過程の特定のフォームが図2に示されている。このいわゆる左右モデルにおいては、状態は左から右へ進む（又はそこに留まる）。このモデルは、信号の性質が時間を経るにつれて変化するモデル化音声について広く用いられる。モデルの状態は、音声を表すものと見ることができる。サブワードユニットについてのモデルにおける状態の数は、例えば5又は6とすることができる。この場合、状態は平均として観察間隔に対応する。図2のモデルにおいては、1つの状態が同一に留まることができる。これは遅い会話に対応させることができる。これに代えて1つの状態をスキップすることができる。これは速い会話に対応させることができる（図2では平均レートの2倍まで）。不連続マルコフ過程の出力は、時間の各時点の状態の集合であり、それぞれの集合が観察できるイベントに対応する。音声認識システムについて、不連続マルコフ過程の

(11)

特表平10-508391

概念は、観察が状態の確率的機能である場合に拡張される。これは、二重確率的処理になる。状態変化の基本的な確率的処理は遮蔽（遮蔽マルコフモデル、HMM）され、観察のシーケンスを生成する確率的処理を通して観察することができるだけである。

音声について、観察は連続的な信号を表す。この観察は、例えば32乃至256のベクトルの有限のアルファベットから選択された不連続の記号に量子化することができる。この場合、このモデルの各状態について不連続的確率密度を用いることができる。量子化に伴う品質低下を避けるため、多くの音声認識システムが連続的確率密度を用いる。一般的に、密度は、ガウス（正規分布）又はラプラス密度のような対数凹面又は楕円対称密度から導出される。学習の間、学習データ（学習観察シーケンス）が初期モデルを用いる状態の中にセグメント化される。これは、各状態について観察の集合を与える。次に、各状態の観察ベクトルをクラスタ化する。システムの複雑度及び学習データの量に応じて、各状態について例えば32乃至120のクラスタがある。各クラスタは自己の例えばガウス密度のような密度を持つ。密度は例えば平均ベクトルのような基準ベクトルで表現される。ここで得られる状態についての観察密度は、クラスタ密度のウェイト付けされた総計である。

音声信号（観察シーケンス）から1つの音声認識ユニット（例えばワード又はサブワードユニット）を認識するため、観察シーケンスを生成する各音声認識ユニットについて尤度を計算する。最大尤度を持つ音声認識ユニットを選択する。より大きい観察シーケンスを認識するため、平均化の方法が用いられる。第1レベルから始めて、尤度を以前の通りに計算する。モデルの最後の状態に到達するとより高いレベルに切換え、残余の観察について同一の処理を繰り返す。最後の観察を処理すると、最大尤度を持つパスを選択し、このパスを逆にたどって、含まれる音声認識ユニットのシーケンスを決定する。

尤度の計算は、観察（特徴）ベクトルとクラスタを表現する各基準ベクトルとの距離の各状態における計算を含む。特に例えば、1状態あたり64個のクラスタを持ち、各サブワードユニット当たり5個の状態を持ち、40個のサブワードユニ

(12)

特表平10-508391

ットを持つ連続的観測密度HMMを用いる大語音音声認識システムにおいては、例えば32次元ベクトル間の12800個の距離計算を含む。これらの計算を各観測について繰り返す。この結果、尤度の計算が計算リソースの50%～75%を消費することもあり得る。各観測ベクトル \vec{o} について密度（及び対応する基準ベクトル）

の部分集合を選択すること及び部分集合について観測ベクトルの尤度を計算することは、E.Bocchieriの"Vector quantization for the efficient computation of continuous density likelihoods", Proceeding of ICASSP, 1993年、692-695頁から既知である。選択した部分集合の一部ではない密度の尤度を近似する。既知の方法によれば、学習する間に全ての密度を隣接部分にクラスタ化する。更に、各隣接部分について1つのコードワードからなるベクトル量子化装置を定義する。各コードワードについて、そのコードワードに隣接する密度の部分集合を定義する。この部分集合の定義は、例えば、予めシステムの学習の間に行っておく。認識の間に、観測ベクトルをコードワードの1つに量子化し、観測の尤度が計算される密度の部分集合としてそのコードワードについて定義された部分集合を用いることにより、予め定義された複数の部分集合から部分集合を選択する。この方法の欠点は、部分集合が与えられた基準ベクトルに基づいて静的に定められることである。特に、予め定められた部分集合の境界近くである観測ベクトルについて、選択された部分集合が実際に、観測ベクトルから隣接する部分集合の基準ベクトルより遠い多くの基準ベクトルを含むことである。従って、パターンエラーレートを小さくするためには、選択される部分集合が比較的大きくなくてはならない。

本発明の目的は、各観測ベクトルについて、更にその観測ベクトルに基づく基準ベクトルの部分集合を選択するために用いられる種類の方法及びシステムを提供することにある。他の目的は、小さいエラーレートのパターンを認識するためのポテンシャルを与える方法及びシステムを提供することにある。更に他の目的は、パターンエラーレートの著しい増加なしに、最大尤度計算のために必要な処理時間の比率を減らすためのポテンシャルを与える方法及びシステムを提供することにある。

(13)

特許平10-508391

この目的を達成するため、本発明による方法は、

各入力観察ベクトル \vec{o} についての基準ベクトルの部分集合 $\{\vec{\mu}_i\}$ の選択が、
前記入力観察ベクトル \vec{o} と前記集合 $\{\vec{\mu}_i\}$ の各基準ベクトルとの間の非類似度
の大きさの計算、及び、基準ベクトルの前記部分集合 $\{\vec{\mu}_i\}$ として前記入力観
察ベクトル \vec{o} との非類似度の大きさが最小である多数の基準ベクトル $\vec{\mu}_i$ の使用

を含むことを特徴とする。類似度に基づき、各観察ベクトルについて動的に基準
ベクトルの部分集合を作ることにより、高い尤度を持ち、観察ベクトルに近い基
準ベクトルを含む部分集合が選択される。これは、更に正確なパターン認識への
道を開く。更に、観察ベクトルに近くないベクトルは、高い尤度を持つ場合でも
この部分集合から除外される。これは、速いパターン認識への道を開く。

EP-A-627-726から、ツリー構造を用いてツリーサーチを行うことにより、基準
ベクトルを組織化することによって、基準ベクトルの部分集合を選択することが
既知であることに注意すべきである。ツリーの最も高いレベルにおいては、根の
ノードは全ての基準ベクトルを表す。このツリーの1段下のレベルにおいては、
複数の中間ノードが、それぞれ基準ベクトルの隣接する部分集合を表す。ここで
は、部分集合が基準ベクトルの全体の集合から集まっている。これは、それ以下の
レベルについても、最も下のレベルでツリーのそれぞれの葉ノードが実際の基準
ベクトルを表示するまで繰り返される。パターン認識の間、各入力観察ベクトル
について、ツリーサーチが根より下の1つのレベルでスタートして遂行される。
このレベルのそれぞれのノードについて尤度が計算され、観察ベクトルがノード
により表示される部分集合によって生成される。このため、各部分集合は、部分
集合ベクトルによって表現され、観察ベクトルがこれと比較される。最大尤度を
持つ1又は複数のノードが選択される。これらのノードについて、それぞれが
基準ベクトルの異なる部分集合を表示し、1つ低いレベルで、同一の処理が繰り
返される。この方法により、最終的に多数の葉ノードが選択される。選択された
葉ノードによって表示される基準ノードは、最終的に選択される基準ベクトルの
部分集合から集まる。この方法を用いることにより、選択される部分集合は実際
の観察ベクトルに基づく。基準ベクトルがツリー構造に予め配置され、ツリーサ
ーチを用いて選択されているので、個々の基準ベクトルは、部分集合の選択の間

(14)

特表平10-508391

に観察ベクトルと比較されることはない。これは、尤度に対して、最も近い基準ベクトルが選択される部分集合の一員であるという逆の作用をする。

本発明による第1のバージョンにおいては、この方法は、前記方法が、各基準ベクトル $\overline{\mu}_i$ を量子化して量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ とし、且つ、基準ベクトルの部分集合 $(\overline{\mu}_i)$ の選択が、各入力観察ベクトル \overline{o} について次のステップ、即ち

入力観察ベクトル \overline{o} を量子化して量子化入力観察ベクトル $\overline{R}(\overline{o})$ とするステ

ップ、

各量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ に対する前記量子化入力観察ベクトル $\overline{R}(\overline{o})$ との距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算するステップ、及び
前記入力観察ベクトル \overline{o} と前記基準ベクトル $\overline{\mu}_i$ との間の前記非類似度の大きさとして、前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を使用するステップ

を含むことを特徴とする方法である。ベクトルを量子化することにより、ベクトルの複雑さが減少し、量子化された観察ベクトルと量子化された基準ベクトルとの間の距離を効率的に計算することができるようになる。この量子化されたベクトル間の距離は、実際のベクトル間の距離の積算と同一と見做することができるので、部分集合を選択するために用いられる。

本発明による他のバージョンにおいては、この方法は、

ベクトル \overline{x} (基準ベクトル $\overline{\mu}_i$ 、又は観察ベクトル \overline{o})を量子化ベクトル $\overline{R}(\overline{x})$ とする量子化が、符号ベクトルの各成分を、ベクトル \overline{x} の対応する成分が負値を持つならば第1二値 b_1 に割当て、ベクトル \overline{x} の対応する成分が正値を持つならば第2二値 b_2 に割当てる方法で、二値への割当てによる符号ベクトル $\overline{S}(\overline{x})$ の計算を含むことを特徴とする。二値を用いることにより、極めて効率的な計算及びマイクロプロセッサを用いる記憶が可能になり、一方、復号ベクトルがベクトルの合理的な近似を与える。

本発明による他のバージョンにおいては、この方法は、

(15)

特許平10-508391

前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ の計算が、ベクトル $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_i)$ とのハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ の計算を含むことを特

徴とする。二値ベクトルのハミング距離は、極めて効率的に計算することができる。

本発明による他のバージョンにおいては、この方法は、

前記量子化が、更に、ベクトル \overline{x} の L_1 ノルムの計算、及び、前記ノルムと前記符号ベクトル $\overline{S}(\overline{x})$ との乗算を含むことを特徴とする。これはベクトルのよい近似を与える。基準ベクトルの L_1 ノルムは予め計算することができ、且つ観

察ベクトル \overline{o} の L_1 ノルムは各観察ベクトルについて1回だけ計算すればよいの

で、追加の計算がオンラインの効率に重大な影響を与えることはない。

本発明による他のバージョンにおいては、この方法は、

前記量子化が、更に、符号ベクトル $\overline{S}(\overline{x})$ のベクトル \overline{x} のディメンションに

よる除算と $1/r$ 乗を含むことを特徴とする。これは、ベクトルの更により近似を与える。

本発明による他のバージョンにおいては、この方法は、

前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ の計算が、

各ベクトル $\overline{\mu}_i$ の L_1 ノルム $\|\overline{\mu}_i\|_1$ の計算、及び

各入力観察ベクトル \overline{o} について、

ベクトル \overline{o} の L_1 ノルム $\|\overline{o}\|_1$ の計算、及び

ベクトル $\overline{S}(\overline{o})$ の各ベクトル $\overline{S}(\overline{\mu}_i)$ に対するハミング距離

$H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ の計算

を含むことを特徴とする。

基準ベクトル $\overline{\mu}_i$ の L_1 ノルム $\|\overline{\mu}_i\|_1$ は予め計算することができ、且つ観

察ベクトル \overline{o} の L_1 ノルム $\|\overline{o}\|_1$ は各観察ベクトルについて1回だけ計算すれ

ばよいので、距離の計算を、極めて効率的に計算することができるハミング距離

(15)

特表平10-508391

$H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_*))$ の基本的な計算に減らすことができる。

本発明による他のバージョンにおいては、この方法は、

ベクトル $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_*)$ との前記ハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_*))$ の計算が、
 $\overline{S}(\overline{o})$ とベクトル $\overline{S}(\overline{\mu}_*)$ との対応する成分の二値 XOR 値を差分ベクトル

の各成分に割当てることにより差分ベクトルを計算すること、

前記差分ベクトルで値 1 を持つ成分の数を計算することにより差分数を求めること、及び

ハミング距離として前記差分数を用いること

を含むことを特徴とする。二値 XOR 値は、多くの成分について 1 回の操作で極めて効率的に計算することができる。

本発明による他のバージョンにおいては、この方法は、前記方法が、

0 又は 1 の二値を持つ成分からなる各 N 次元ベクトルについて、値 1 を持つ成分の数を表す対応する数を特定するテーブルを作ること、及び

前記差分ベクトルを前記テーブルに位置させ且つ前記対応する数をハミング距離として用いて前記差分数を求めること

を含むことを特徴とする。ベクトル中の 1 の要素の数を予め数え、これをテーブルに格納することにより、効率を更に高めることができる。

本発明による他のバージョンにおいては、この方法は、前記方法が、

予め定められた入力観察ベクトル \overline{o} についての基準ベクトルの部分集合を選択

し、続く多数の観察ベクトルに対して同一の部分集合を用いるようにすることを特徴とする。多数の以後の観察ベクトルに対して同一の部分集合を用いることにより、効率が更に改善される。

本発明による他のバージョンにおいては、この方法は、前記方法が、

(17)

特表平10-508391

1つの入力観察ベクトル \overline{o} について基準ベクトルの前記部分集合を選択し、表

現されない各基準ユニットについて、代表基準ベクトルを前記部分集合に加算することにより、各基準ユニットが前記部分集合の少なくとも1つの基準ベクトルで表現されるようにすることを特徴とする。各基準ユニットが部分集合中に表現されていると、認識の精度が改善される。

本発明による他のバージョンにおいては、この方法は、前記方法が、

各基準ユニットが前記部分集合の少なくとも1つの基準ベクトルで表現されるようにし、代表基準ベクトルとして前記基準ユニットを表現し且つ前記入力観察ベクトル \overline{o} に対して最小距離を持つ部分集合から基準ベクトルを選択することに

より、各基準ユニットについて前記代表基準ベクトルを選択することを特徴とする。観察ベクトルは徐々に変化する傾向があるので、特定の観察ベクトルに対する基準ユニットの最良の表現であることが見出された基準ベクトルは、以後の観察ベクトルについての部分集合を表現するために良い候補である。

本発明の目的を達成するため、本発明によるシステムは、

各入力観察ベクトル \overline{o} についての基準ベクトルの部分集合 $\{\overline{\mu}_i\}$ の選択が、前記入力観察ベクトル \overline{o} と前記集合 $\{\overline{\mu}_i\}$ の各基準ベクトルとの間の非類似度の大きさの計算、及び、基準ベクトルの前記部分集合 $\{\overline{\mu}_i\}$ として前記入力観

察ベクトル \overline{o} との非類似度の大きさが最小である多数の基準ベクトル $\overline{\mu}_i$ の使用

を含むことを特徴とする。

本発明によるシステムの第1の実施例は、

前記基準パターンデータベースが、更に、各基準ベクトル $\overline{\mu}_i$ について量子化基準ベクトル $R(\overline{\mu}_i)$ を格納し、及び

基準ベクトルの部分集合 $\{\overline{\mu}_i\}$ の選択が、各入力観察ベクトル \overline{o} について次

のステップ、即ち

(18)

特表平10-508391

入力観察ベクトル \overline{o} を量子化して量子化入力観察ベクトル $\overline{R}(\overline{o})$ とするステ

ップ、

各量子化基準ベクトル $\overline{R}(\overline{\mu}_i)$ に対する前記量子化入力観察ベクトル $\overline{R}(\overline{o})$ の距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算するステップ、及び

前記入力観察ベクトル \overline{o} と前記基準ベクトル $\overline{\mu}_i$ との間の前記非類似度の大きさとして、前記距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を使用するステップ

を含むことを特徴とする。

本発明によるシステムの他の実施例は、

各基準ベクトル $\overline{\mu}_i$ について基準パターンデータベースが基準ベクトル $\overline{\mu}_i$ の L_i ノルム $\|\overline{\mu}_i\|$ を含み、

前記局所化装置が、各入力観察ベクトル \overline{o} について、

ベクトル \overline{o} の L_i ノルム $\|\overline{o}\|$ 、及び、ベクトル $\overline{S}(\overline{o})$ の各ベクトル $\overline{S}(\overline{\mu}_i)$ に対するハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ を計算し、及び

L_i ノルム $\|\overline{o}\|$ 及びハミング距離 $H(\overline{S}(\overline{o}), \overline{S}(\overline{\mu}_i))$ を、前記基準パターンデータベースに格納された L_i ノルム $\|\overline{\mu}_i\|$ と組合せて、距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算する

ことを特徴とする。

本発明によるシステムの他の実施例は、

0又は1の二値を持つ成分からなる各N次元ベクトルについて、値1を持つ成分の数を表す対応する数を特定するテーブルを格納するメモリーを具え、及び前記差分ベクトルを前記テーブルに位置させ且つ前記対応する数をハミング距離

として用いて前記差分数を求めること

を含むことを特徴とする。

これら及び他の本発明の観点は、次の図面に示された実施例を参照することにより、明瞭になり且つ明らかにされる。

図1は連続的音声認識に通常用いられる処理ステップを説明する図、

図2は左右不連続マルコフ過程の例を示す図、

(19)

特表平10-508391

図3は本発明によるシステムの一実施例のブロック図、

図4は第1ベクトル量子化の二次元表示を示す図、

図5は第1ベクトル量子化によって得られた結果を示す図、

図6はベクトル成分の量子化を説明する図、

図7は第2ベクトル量子化の二次元表示を示す図、

図8は第2ベクトル量子化によって得られた結果を示す図である。

図3は、本発明による、音声又は画像のような連続的な物理量から導出される時系列入力パターン20を認識するシステム10のブロック図を示す。入力手段30は周期的に物理量にアクセスする。音声については、これは通常、例えば6.67kHz又は16kHzのような正規の間隔で行われる物理量のサンプリング及びこのサンプルのデジタル化を含む。入力手段30は、例えば32msecに対応する音声信号の連続的なサンプルのグループを処理し、特徴の代表ベクトル（入力観察ベクトル \overline{r} ）を与える。このようにして、入力パターンを表す入力観察ベクトルの時間シーケンスが発生される。典型的には、入力手段30はマイクロホン、アナログ／デジタル変換器、及び例えばデジタル信号プロセッサ（DSP）のようなプロセッサを用いて実現される。付加的に、入力手段30が、音声が実際に受信された時だけサンプリングを行う音声検出器を具えてもよい。入力信号をサンプリングしデジタル化する代わりに、信号をデジタル形式でメモリーに格納しておくか又は通信ネットワークを経てデジタルで供給するようにしてもよい。基準パターンを格納するために基準パターンデータベース40が用いられる。前述したように、音声認識ユニットが音声を認識するための基準パターンとして用いられる。各基準パターンは、基準ユニットのシーケンスを含む。各基準ユニットは、少なくとも1つの対応する基準ベクトル $\overline{\mu}$ によって表現される。全ての基準ベクトル

ルが集まって基準ベクトルの集合 $\{\overline{\mu}_i\}$ を形成する。遮蔽マルコフモデルに基

づくパターン認識を用いると、各基準パターンは遮蔽マルコフモデルによってモデル化され、このモデルの状態は基準ユニットに対応する。例えばガウス密度又はラプラス密度のような連続観察密度を用いると、基準ベクトルは密度の平均ベ

(20)

特許10-508391

クトルに対応する。基準データベース40は、例えばハードディスク、ROM又はRAMのようなメモリーに統合データベース又は個別データファイルとして格納されていてもよい。

システム10は、入力パターンに対応する基準パターンをデータベース40中で局所化するために、更に局所化装置50を具える。局所化装置50は、DSP又はマイクロプロセッサを用いて実現される。局所化された基準パターンは、認識された基準パターンとして参照される。前述のように、本発明は観察ベクトルの尤度の計算を含む。各連蔽マルコフモデル及び各モデルの状態 s について、観察ベ

クトル \vec{o} の尤度は

$$p(\vec{o}) = \sum_{k=1}^N w(k) \cdot p(\vec{o}|k)$$

で与えられる。ここで、 $w(k)$ は k 番目の観察混合密度（クラス）の重み、 N は1つの状態についてのクラスの数である。簡単にするため、状態 s は式中に示されない。音声認識システムは、クラスタの確率分布を作るために、通常、ラプラス又はガウス確率密度を用いる。L₁ノルムを用いて、

$$d_1(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|_1 = \left(\sum_{i=1}^D |x_i - y_i| \right)$$

のように定められる。ここでL₁ノルムはラプラス密度のために用いられ、L₂ノルムはガウス密度のために用いられる。確率について可能な式の1つとして、

$$p(\vec{o}) = \sum_{k=1}^N w(k) \cdot a \cdot e^{-b \|\vec{o} - \vec{\mu}(k)\|_1}$$

が与えられる。ここで、基準ベクトル $\vec{\mu}(k)$ は k 番目の観察混合密度の平均ベクトルである。係数 a 及び b は、観察ベクトル \vec{o} が全ての可能な値を満たしている場合に、確率が1まで積分されるようにするためのものである。この式の種々の形又は拡張は既知である。例えば、次の3つのガウス密度が与えられる。

(21)

特表平10-508391

$$\begin{aligned}
 \text{完全共分散行列} \quad K_{xx}: p(\bar{o} | k) &= \frac{1}{\sqrt{(2\pi)^D \det K_{xx}}} \cdot e^{-\frac{1}{2}(\bar{o}-\bar{\mu})^T K_{xx}^{-1}(\bar{o}-\bar{\mu})} \\
 \text{対角線共分散行列} \quad (K_{xx})_{dd} = \sigma_d^2: p(\bar{o} | k) &= \frac{1}{\sqrt{(2\pi)^D \prod_{d=1}^D \sigma_d^2}} \cdot e^{-\frac{1}{2} \sum_{d=1}^D \frac{(\bar{o}_d - \bar{\mu}_d)^2}{\sigma_d^2}} \\
 \text{スカラー共分散} \quad K_{xx} = I \cdot \sigma_{xx}^2: p(\bar{o} | k) &= \frac{1}{\sqrt{(2\pi)^D (\sigma_{xx}^2)^D}} \cdot e^{-\frac{1}{2\sigma_{xx}^2} \|\bar{o} - \bar{\mu}\|^2}
 \end{aligned}$$

L_r ノルム以外の他の距離尺度を用いることができることは明らかである。

尤度計算を行う前に観察ベクトル \bar{o} 及び平均ベクトル $\bar{\mu}(k)$ の尺度構成を行う

ことが望ましい。尺度構成は、その項がプロセッサの精度範囲以下に落ちることを防ぐため及び密度の分散に従ってベクトルを規格化するために用いることができる。この尺度構成は、ベクトルと対角線 D X D マトリクス V との乗算によって遂行される。ここで、D は特徴ベクトル空間のディメンションである。

$$V = \begin{bmatrix} v_1 & 0 & \dots & \dots & 0 \\ 0 & v_2 & \dots & \dots & \dots \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots & v_{D-1} & 0 \\ 0 & \dots & \dots & 0 & 0 & v_D \end{bmatrix}$$

マトリクス要素 V_1 乃至 V_D は学習の間に積算される。基準ベクトルの全集合からスタートし、各成分について個別に、共通分散を計算し、1つのプール化分散ベクトルを得る。対応するプール化標準偏差によって全ての基準ベクトルからベクトル成分が分けられ、その成分を必要な範囲に入れるために値 γ で再び尺度構成を行う。値 γ は、各ベクトル成分に対して同一である。尺度構成の他の形が既知である。基準ベクトルが予め尺度構成され、実際の尤度の計算のスタートの前に、観察ベクトルがただ一度だけ尺度構成されることが望ましい。

密度の性質により、確率の総計は最大値即ち最大確率に寄与する密度に近似することができる。これは、入力パターンに対応する基準パターンを局所化するキ

(22)

特表平10-508391

ステップが、観察ベクトルに最も近い基準ベクトルを見出すこと（最近接近サーチ）であることを意味する。

$$p(\hat{o}) = \max\{w(k) \cdot a \cdot e^{-\frac{1}{2}\|\hat{o}-\bar{\mu}(k)\|^2} \mid k = 1, \dots, N\}$$

対数をとることにより、これは、

$$\log(p(\hat{o})) = -\min\{b\|\hat{o}-\bar{\mu}(k)\|^2 - \log(w(k)) \mid k = 1, \dots, N\} - \log(a)$$

を与える。定数 $\log(a)$ は無視できる。項 $\log(w(k))$ を個別に減算する代わりに、

$$\begin{aligned}\bar{p}^T &= (b\bar{o}^T, 0), \\ \bar{q}(k)^T &= (b\bar{\mu}^T, (-\log(w(k)))^T)\end{aligned}$$

で定義される新たに拡張されたベクトル \bar{p} 及び $\bar{q}(k)$ を導入してもよい。この式

においては、 $-\log(w(k)) > 0$ であることに注意すべきである。拡張されたベクトルを用いると、

$$\log(p(\hat{o})) = -\min\{\|\bar{p}-\bar{q}(k)\|^2 \mid k = 1, \dots, N\} - \log(a)$$

が与えられる。 \bar{p} 及び $\bar{q}(k)$ が更に1つの成分を持つので、 D を特徴ベクトル空間

間の次元とすると、それらの次元は $D+1$ である。

以下においては、ベクトル \bar{x} 及び \bar{y} が参照される。ここで、 $\bar{x} = \bar{o}$ 及び $\bar{y} = \bar{\mu}$ である。次に説明するように、 $\bar{x} = \bar{p}$ 及び $\bar{y} = \bar{q}$ と読むことにより、ベクトル \bar{p} 及び \bar{q} に同一の概念を適用できることが明らかである。

従来のパターン認識は、例えば最大尤度又はその一部分として入力観察ベクトル \bar{o} と各基準ベクトルとの間の距離のような、ベクトル類似度の計算を含む。前述したように、大語彙システムについての最大尤度の計算は、例えば32次元ベクトルの間の10,000個の距離の計算を含む。これらの計算は、各観察について繰り返される。全体の全ての計算を行う代わりに、図3の局所化装置50が最初に基準ベクトルの全集合 $\{\bar{\mu}_i\}$ から基準ベクトルの部分集合 $\{\bar{\mu}_j\}$ を選択する。局所化装置50は次に部分集合 $\{\bar{\mu}_j\}$ の基準ベクトルについての全ての距離のみを

(23)

特表平10-508391

計算する。本発明によれば、局所化装置50が、観察ベクトル \vec{o} と全集合 $\{\vec{\mu}_i\}$ の各基準ベクトルとの類似度の尺度を計算して部分集合 $\{\vec{\mu}_i\}$ を決定する。次

に、類似度の最小尺度を計算してある多数の基準ベクトルを部分集合として用いる。その数は、固定数又は基準ベクトルの全数に対する割合であってもよい。これに代えて、局所化装置50が、距離 $d(\vec{R}(\vec{o}), \vec{R}(\vec{\mu}_i))$ が予め定められたしきい値 T 以下の基準ベクトル $\vec{\mu}_i$ を選択する。最適の数又は、性能のかなり

の向上を得てしかも認識の精度に著しく影響し過ぎないように十分に低いしきい値を決定するために、テストを行うことが好ましい。類似度の尺度は、観察ベクトルと比較して類似度（「距離」）に関する基準ベクトルの順位付けを行うために用いられる。この順位付けは、高い尤度で部分集合が観察ベクトルに近い基準ベクトルを含む限り、パターン認識のために用いられる実際の距離尺度（尤度計算）を用いて完全に正確を期す必要はない。これにより、実際の全ての距離尺度に課される制約に比べて、類似度の尺度に対しては制約が少なくなる可能性がある。全ての距離の計算に比べて、ずっと効率がよい近似距離の計算を可能にするように、類似度の尺度を選択することが好ましい。

本発明による他の実施例においては、局所化装置50が、以下によって部分集合 $\{\vec{\mu}_i\}$ を決定する。

- ・各基準ベクトル $\vec{\mu}_i$ を量子化して量子化基準ベクトル $\vec{R}(\vec{\mu}_i)$ とすること。

これは、基準ベクトルが決定される時に、学習の間に前もって行うことができる。この場合、実際の音声認識が開始される前に、量子化ベクトル $\vec{R}(\vec{\mu}_i)$ を

基準パターンデータベース40に格納できることが好ましい。これは、このシステムのオンライン性能を向上させる。

(24)

特表平10-508391

・各入力観察ベクトル \overline{o} について:

－入力観察ベクトル \overline{o} を量子化して量子化ベクトル $\overline{R}(\overline{o})$ とすること、

－量子化ベクトル $\overline{R}(\overline{o})$ の各量子化ベクトル $\overline{R}(\overline{\mu}_i)$ に対する距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ を計算すること、及び

－部分集合 $\{\overline{\mu}_i\}$ として、全集合 $\{\overline{\mu}_i\}$ 中の、距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ が最小である多数の基準ベクトル $\overline{\mu}_i$ を用いること。このようにして、距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}_i))$ が観察ベクトル \overline{o} と基準ベクトル $\overline{\mu}_i$ との間の

類似度の尺度として用いられる。

ベクトルの量子化はベクトルの複雑さを減少させる。次の実施例においては、量子化の種々の形を説明し、量子化ベクトル相互間の簡潔な距離計算を示す。1つの方法は、例えば量子化ベクトルとして二値成分を持つベクトルに比例するベクトルを用いることにより、ベクトル成分の複雑さを減少させることである。他の方法は、元のベクトルの成分と同様の複雑さを持つベクトル成分が少ない（低次元）量子化ベクトルを用いることである。量子化ベクトルとして、幾つかの成分が複雑さを低減した成分であり、残余の成分が元のベクトルと同様の複雑さを持つ成分であるベクトルを用いることによって、両方法を組合せてもよい。

他の実施例においては、ベクトル \overline{z} の量子化ベクトルとして、符号ベクトル $\overline{S}(\overline{z})$ が用いられる。同じ量子化が \overline{x} 及び \overline{y} 両者に対しても用いられる。 $\overline{S}(\overline{z})$ の成分は、 \overline{z} の対応する成分の符号を表す二値を持つ。

$$\overline{z} \rightarrow \overline{S}(\overline{z})$$

ここで、 $\overline{S}(\overline{z}) = (\text{符号}(z_1), \dots, \text{符号}(z_D))^T$ 、及び

$$z_i < 0 \quad \text{ならば} \quad \text{符号}(z_i) = b_1$$

$$z_i > 0 \quad \text{ならば} \quad = b_2$$

である。ここで $i = 1, \dots, D$ である。

\overline{z} の対応する成分が値0を持つ $\overline{S}(\overline{z})$ の成分については、 b_1 又は b_2 のい

ずれかを選択することができる。全ての量子化ベクトルについて同一の選択を行うことが好ましい。これに代えて、成分の0値を表すために第3の独立値を用いてもよい。この定義は、元のベクトル成分が例えば符号付き2バイト整数のよう

(25)

特表平10-508391

な符号付きの値を用いて表されるシステムに対して与えられている。下半分の値 (0..32767) を b_1 に割当て、上半分の値 (32768..65535) を b_2 に割当てることにより、同一の概念を符号なしの値に適用できることが理解されるであろう。

この方法のよい説明が、 b_1 に対して -1 を、 b_2 に対して $+1$ を用いて達成される。

$$\overline{z} \rightarrow \overline{S}(\overline{z})$$

ここで、 $\overline{S}(\overline{z}) = (\text{符号}(Z_1), \dots, \text{符号}(Z_D))^T$ 、及び

$$z_i < 0 \quad \text{ならば} \quad \text{符号}(Z_i) = -1$$

$$z_i < 0 \quad \text{ならば} \quad \text{符号}(Z_i) = -1$$

$$z_i \geq 0 \quad \text{ならば} \quad = 1$$

である。ここで $i = 1, \dots, D$ である。

この場合、0 成分も $+1$ によって表される。幾何学的には、ベクトル自身と同一の扇形空間の中心で、「単位」ベクトル上に各ベクトルを投影したように見える。図4は、ベクトル $(3,1)^T$ がベクトル $(1,1)^T$ 上に投影される2次元の例を示す。同じことが扇形空間(象限) I において全てのベクトルに適用される。同様に、扇形空間 II のベクトルは $(-1,1)^T$ 上に投影され、扇形空間 III のベクトルは $(-1,-1)^T$ 上に投影され、扇形空間 IV のベクトルは $(1,-1)^T$ 上に投影される。例

として、3次元特徴ベクトル空間において、ベクトル $\overline{x} = (51, 72, 46)^T$ は、次の

ように量子化される。

$$\overline{x} = (51, -72, 46)^T \rightarrow (1, -1, 1)^T = \overline{S}(\overline{x})$$

ベクトル \overline{x} と \overline{y} との間の距離 $d_r(\overline{x}, \overline{y})$ が、それぞれの符号ベクトル間で距離 $d_r(\overline{S}(\overline{x}), \overline{S}(\overline{y}))$ により置き換えられる。

$$d_r(\overline{x}, \overline{y}) \rightarrow d_r(\overline{S}(\overline{x}), \overline{S}(\overline{y}))$$

図5は、上記で与えられたベクトル \overline{x} 及び10個の平均ベクトル(プロトタイプ)を表現する10個の拡張された三次元ベクトル $\overline{y}_1, \dots, \overline{y}_{10}$ についてのこの方

法を説明する図である。この例においては、ベクトル要素は -99 から $+99$ までの範囲でランダムに選択された整数である。音声認識システムにおいては、ベクトル

(26)

特表平10-508391

ル成分は典型的には連続値（数学的観点：Rからの要素）であり、コンピュータ上では通常2又は4バイト整数又は浮動数として表されることに注意すべきである。図の各欄は、10個のベクトル \overline{y}_i のインデックス $i = 0, \dots, 9$ 、ベクトル \overline{y}_i 、 \overline{x} と \overline{y} との差分、ベクトル間の距離 d 、 $(\overline{x}, \overline{y}_i)$ 、ベクトル \overline{y}_i の符号、符号ベクトル $\overline{S}(\overline{x})$ と $\overline{S}(\overline{y}_i)$ との差分及び符号ベクトル間の距離 \overline{d} 、 $(\overline{S}(\overline{x}), \overline{S}(\overline{y}_i))$ を示す。全距離計算（L₁ノルム）を用いると、距離は昇順に $\overline{y}_8, \overline{y}_7, \overline{y}_1, \overline{y}_9, \overline{y}_4, \overline{y}_0, \overline{y}_5, \overline{y}_3, \overline{y}_6$ 、及び \overline{y}_2 である。近似距離を用いると次の4つのグループに分かれる（昇順に）。即ち \overline{y}_8 、及び \overline{y}_7 、次に \overline{y}_0 、 \overline{y}_1 、及び \overline{y}_9 、次に $\overline{y}_5, \overline{y}_3, \overline{y}_4, \overline{y}_6$ 、及び \overline{y}_2 である。既に、三次元ではベクトル

は比較的良好にソートされている。符号ベクトル成分の平均して半分が値1を持つ（即ち、符号ベクトルが空間によく分布しており、扇形空間に過剰のクラスター

がない）と、実用上の良好な結果が得られることが明らかである。これは、各基準ユニット（混合）について独立に近似を実行して達成することが好ましい。各基準ユニットについて、混合平均ベクトルを計算し、基準ユニットを表現する全ての基準ベクトルから差し引く。次に、基準符号ベクトルを計算する。観察ベクトルの符号ベクトルを計算する前に、基準ユニットの混合平均ベクトルを観察ベクトルから差し引く。次に、観察符号ベクトルを計算し、観察符号ベクトルをその基準ユニットについて全ての基準符号ベクトルと比較する。この操作は各基準ユニットに対して独立に実行される。この操作は、各基準ユニット（それを表現する全ての基準ベクトル及び観察ベクトル）について座標系を移動すると見ることができる。この方法は、大語彙連続音声認識のためのフィリップス研究システムでテスト済である。各基準ユニットについて、32個の基準ベクトルを用いた場合、実験によれば、正確な距離計算を達成するために基準ベクトルのほぼ20%の部分集合を選択すれば充分であることが示された。これは、距離計算に必要なランタイムを50%減らした。

(27)

特表平10-508391

他の実施例においては、ベクトル $\overline{S}(\overline{x})$ と $\overline{S}(\overline{y})$ とのハミング距離 $H(\overline{S}(\overline{x}), \overline{S}(\overline{y}))$ を計算することにより、ベクトル $\overline{S}(\overline{x})$ と $\overline{S}(\overline{y})$ との距離 $d(\overline{S}(\overline{x}), \overline{S}(\overline{y}))$ を計算する。ハミング距離は極めて高速で計算でき

る。符号ベクトルは、0及び1の要素のみのベクトルで表現されることが望ましい。例えば、八次元符号ベクトル $(1, -1, -1, -1, 1, 1, -1, 1)^T$ は $(1, 0, 0, 0, 1, 1, 0, 1)^T$ で表現される。これらのベクトルは、ビットのシーケンスとしてコンピュータのメモリーに格納されることが望ましい。ベクトルは、例えば、バイト、16ビット又は32ビットワードのような、コンピュータの望ましい配列に基づいて配列されることが好ましい。ハミング距離の計算の次のステップとして、全ての2つのベクトルについてXOR関数を計算し、差分ベクトルを算出する。差分ベクトルの各成分は、 $\overline{S}(\overline{x})$ 及び $\overline{S}(\overline{y})$ の対応する成分の二値XOR値を含む。大

部分のマイクロプロセッサにおいて、全てのコンピュータワードについてXOR関数の計算が1つの動作で行われる。例えば、 \overline{x} が $(1, 0, 0, 0, 1, 1, 0, 1)^T$ で表され、 \overline{y} が $(1, 0, 1, 0, 0, 0, 1, 1)^T$ で表される場合、これは、距離ベクトル $(0, 0, 1, 0, 1, 1, 1, 0)^T$ を与える。原理的に、ハミング距離はここでは、差分ベクトルの1

の要素の数を数えることによって計算される。この例においては、ハミング距離は4である。必要ならばハミング距離を2倍すると、図5で用いた距離と同一の距離を与える。

他の実施例においては、各差分ベクトルについてのハミング距離が前もって計算され、テーブルの形で図3のメモリー60に格納される。例として、ビットシーケンス $(0, 0, 1, 0, 1, 1, 1, 0) = 2E(16進)$ についてはハミング距離4がテーブルの第46(=2E(16進))エントリーに格納される。ベクトル空間のディメンションが大き過ぎる場合(例えば16を超える場合)には、テーブル全体に格納する代りに、ベクトルを例えば16ビットより小さいユニットに分割することが望ましい。各ユニットについて、ユニットのハミング距離を求めるためにテーブルが用いられる。各個別のユニットのハミング距離を合計することによって、完全なハミン

(28)

特表平10-508391

グ距離を得ることができる。

例えば幾つかの成分を除外することにより、符号ベクトルのディメンションを減らすことによって、性能についてこれ以上の改善を達成できることが明らかである。例えば、半分の成分だけ（例えば第1のD/2成分）を用い、これを補償するために得られた積算距離を2倍する。成分の数を処理に適したサイズに減らすことが好ましい。例えば、多くのマイクロプロセッサについては、35次元ベクトルを、32ビットコンピュータワードを形成する1ビット成分を持つ32次元ベクトルに減らすことが望ましい。

ベクトル成分を2つの領域（二値表現）の内の1つに割当ててこれを量子化する代わりに、同様に2を超える領域を用いてもよいことが理解されるであろう。例えば、図6に示すようにベクトルの1つの成分（x1成分）について3つの領域を用いることができる。量子化されたベクトル $\overline{R}(\overline{z})$ の成分は、 \overline{z} の対応す

る成分を表す2ビット値を持つ。

$$\overline{z} \rightarrow \overline{R}(\overline{z})$$

ここで、 $\overline{R}(\overline{z}) = (f(z_1), \dots, f(z_D))^T$ 、及び

$$z_i < a_i \text{ ならば } f(z_i) = '00'$$

$$a_i \leq z_i \leq b_i \text{ ならば } f(z_i) = '01'$$

$$z_i > b_i \text{ ならば } f(z_i) = '11'$$

である。ここで $i = 1, \dots, D$ である。

図6に示すように、前記と同様に、部分集合を求めることによって、二値XOR操作及びハミング距離計算は、量子化されたベクトル間の距離の計算及び実際のベクトルの距離の表示に用いることができる。

代替実施例においては、符号ベクトルに加えて更にベクトルのノルムを用い、

更に正確な近似を与える。この近似は、ベクトル \overline{z} を

$$\overline{z} = \mathcal{I}(\overline{z}) \frac{\|\overline{z}\|}{D} + \Delta \overline{z}$$

(29)

特許平10-508391

と書くと、第1項が \bar{x} についてのよい近似であるという洞察に基づいている。幾

何学的には、これは各ベクトルを回転することによりそのベクトル自身と同一の扇形空間の対角線上にそのベクトルを投影したように見える。図7は L_2 ノルムを用いる二次元の例を示しており、ここではベクトル $(3,1)^T$ がベクトル $(\sqrt{5}, \sqrt{5})^T$ 上に投影されている。

三次元の例においては、ベクトルが $\bar{x}^T = (51, -72, 46)^T$ であり、 L_1 ノルム

を用いると、

$$\bar{x}^T = (51, -72, 46)^T = (1, -1, 1)^T \cdot 56\frac{1}{2} + \Delta \bar{x} = (56\frac{1}{2}, -56\frac{1}{2}, 56\frac{1}{2})^T + \Delta \bar{x},$$

ここで $\Delta \bar{x} = (-5\frac{1}{2}, -15\frac{1}{2}, -10\frac{1}{2})^T$ である。

\bar{x} と \bar{y} との距離については、これにより

$$\|\bar{x} - \bar{y}\|_1 = d_{xy} + \epsilon,$$

である。ここで

$$d_{xy} = \|\mathcal{S}(\bar{x}) \frac{\|\bar{x}\|_1}{D^{\frac{1}{2}}} - \mathcal{S}(\bar{y}) \frac{\|\bar{y}\|_1}{D^{\frac{1}{2}}}\|_1,$$

である。三角不等式は次のようになる。

$$\epsilon \leq \|\Delta \bar{x}\|_1 + \|\Delta \bar{y}\|_1$$

エラー ϵ を無視すると、 \bar{x} と \bar{y} との間の距離 $d(\bar{x}, \bar{y})$ は d_{xy} によって近似される。図8は、上記で与えられたベクトル \bar{x} 及び10個の平均ベクトル(プロトタイプ)を表す10個の拡張された三次元ベクトル $\bar{y}_1, \dots, \bar{y}_{10}$ についてのこの方

法を説明する図である。図の最後の2つの欄は、 L_1 ノルムを用いて、近似距離

d_{xy} 及びベクトル \bar{x} と \bar{y}_i との間の全距離を示している。全距離計算を用いると、距離は昇順に $\bar{y}_8, \bar{y}_1, \bar{y}_9, \bar{y}_0, \bar{y}_4, \bar{y}_{10}, \bar{y}_6, \bar{y}_2, \bar{y}_5$ 及び \bar{y}_3 である。近似距離を用いると、距離は昇順に $\bar{y}_8, \bar{y}_1, \bar{y}_0, \bar{y}_4, \bar{y}_9, \bar{y}_6, \bar{y}_2, \bar{y}_5$ 及び \bar{y}_3 であ

る。既にこのベクトルの三次元ではベクトルのよいソートが達成されている。

(30)

特表平10-508391

この方法は、大語彙連続音声認識のためのフィリップス研究システムでテスト済である。32個の観察密度成分（クラスタ）を用いて次の結果が得られた。

	計算した距離の数	動作時間（％）	ワードエラー（％）
全距離	32	1.00	24.52
HVQ-1	23.9	0.75	24.52
HVQ-2	21.3	0.69	24.52
HVQ-3	18.6	0.62	24.52
HVQ-4	15.9	0.53	25.07
HVQ-5	13.3	0.48	25.07
HVQ-6	10.6	0.42	25.07
HVQ-7	7.9	0.34	25.07
HVQ-8	5.3	0.29	25.34

この表においては、結果の第1行目に、全ての距離を完全に計算する通常の方法の結果を示す。それ以下の8行には、上記の方法（HVQと記す）を用いて、それぞれの場合に異なる数／完全な場合に対する距離の比率を計算した結果を示す。表に示すように、完全な場合の距離のほぼ20%を計算するだけで、最近接サーチのための計算時間の約30%しか必要としないにもかかわらず、ワードエラーの割合は全ての距離を完全に計算する場合に殆ど近くなっている。最近接サーチは全処理時間の50-75%を占めるので、これによる全処理時間の減少は約35-50

%である。

これらの結果を得るために、他の実施例について説明した $d_{\overline{z}}$ を計算するための他の最適化を用いた。ベクトル \overline{z} の近似のために、更に

$$\overline{z} \rightarrow \overline{S}(\overline{z}), \|\overline{z}\|,$$

(31)

特表平10-508391

を用いることができることが明らかである。この近似は前述の近似より正確さが劣るが、差異はベクトルの順序付けに影響しない定数 $1/D^{1/r}$ を乗算することである。

L_r ノルムを用いて近似距離 d_{xy} を

$$d_{xy} = \left(\sum_{i=1}^D \frac{1}{D} |\overline{S}_i(\overline{x})\|\overline{x}\|_r - \overline{S}_i(\overline{y})\|\overline{y}\|_r|^r \right)^{1/r}$$

と書くことができる。 $|\overline{S}_i(\overline{x})| = 1$ である場合、

$|\overline{S}_i(\overline{x})\|\overline{x}\|_r - \overline{S}_i(\overline{y})\|\overline{y}\|_r|$ は、

$\overline{S}_i(\overline{x}) = \overline{S}_i(\overline{y})$ ならば $= |\overline{S}_i(\overline{x})| \cdot |\|\overline{x}\|_r - \|\overline{y}\|_r|$ 、

$\overline{S}_i(\overline{x}) = -\overline{S}_i(\overline{y})$ ならば $= |\overline{S}_i(\overline{x})| \cdot |\|\overline{x}\|_r + \|\overline{y}\|_r|$

であると考え、これは、

$$(d_{xy})^r = \sum_{i \in \mathcal{A}(\overline{x}, \overline{y})} \frac{1}{D} |\|\overline{x}\|_r - \|\overline{y}\|_r|^r + \sum_{i \in \mathcal{B}(\overline{x}, \overline{y})} \frac{1}{D} |\|\overline{x}\|_r + \|\overline{y}\|_r|^r$$

を与える。ここで

$$q_{xy} = \sum_{i \in \mathcal{A}(\overline{x}, \overline{y})} \frac{1}{D}$$

のように定義する。但し、 $i = 1, \dots, D$ である。

ベクトルのディメンションは D であるから、与えられた定義により、 $D q_{xy}$ 個

の要素が同一の符号を持ち、従って $D - D q_{xy} = D(1 - q_{xy})$ 個の要素が別の

符号を持つ。これは、

$$\begin{aligned} \text{一般: } (d_{xy})^r &= q_{xy} |\|\overline{x}\|_r - \|\overline{y}\|_r|^r + \\ &\quad (1 - q_{xy}) |\|\overline{x}\|_r + \|\overline{y}\|_r|^r \end{aligned}$$

を与える。ラプラス密度 (L_1 ノルム) については、距離についてこれは次の近似を与える。

(32)

特表平10-508391

$$\begin{aligned} \text{ラブラシアン: } \|\bar{x} - \bar{y}\|_1 &\approx \\ &\|\bar{x}\|_1 + \|\bar{y}\|_1 - 2 \min(\|\bar{x}\|_1, \|\bar{y}\|_1) \end{aligned}$$

ガウス密度 (L_2 ノルム) については、距離についてこれは次の近似を与える

$$\begin{aligned} \text{ガウシアン: } \|\bar{x} - \bar{y}\|_2^2 &\approx \\ &\|\bar{x}\|_2^2 + \|\bar{y}\|_2^2 - 2(2q_{xy} - 1) \|\bar{x}\|_2 \|\bar{y}\|_2 \end{aligned}$$

この結果、 q_{xy} 並びにベクトル \bar{x} 及び \bar{y} のノルムを用いるだけで距離を積算することができる。ベクトル \bar{y} のノルムは予め計算でき、且つ好ましいことに、図

3の基準パターンデータベース40に格納することができることは明らかである。

更にベクトル \bar{x} のノルムは一度に計算することができ、各ベクトル \bar{y} の計算に利用することができる。 d_{xy} の計算については、 $D = D(q_{xy})$ が \bar{x} 及び \bar{y} の符号ベクトルのハミング距離であることに注意すべきである。 \bar{x} 及び \bar{y} の符号ベクトルのハミング距離として h_{xy} を定義することにより、前述の3つの式は次のように書くことができる。

$$\text{一般: } (d_{xy}) = (1 - \frac{h_{xy}}{D}) \|\bar{x}\|_1 + \|\bar{y}\|_1 - \frac{h_{xy}}{D} \|\bar{x}\|_1 + \|\bar{y}\|_1$$

$$\text{ラブラシアン: } \|\bar{x} - \bar{y}\|_1 \approx \|\bar{x}\|_1 + \|\bar{y}\|_1 - 2(1 - \frac{h_{xy}}{D}) \min(\|\bar{x}\|_1, \|\bar{y}\|_1)$$

$$\text{ガウシアン: } \|\bar{x} - \bar{y}\|_2^2 \approx \|\bar{x}\|_2^2 + \|\bar{y}\|_2^2 - 2(1 - 2\frac{h_{xy}}{D}) \|\bar{x}\|_2 \|\bar{y}\|_2$$

この結果、ハミング距離 h_{xy} 並びにベクトル \bar{x} 及び \bar{y} のノルムを用いて距離を

積算することができる。部分ベクトルに対する操作を行うことにより、精度を改善できたことは明らかである。ベクトル \bar{x} 及び \bar{y} は N 個の部分ベクトルに分解されて、部分ノルムが計算される。ベクトル \bar{x} について、これは次のように定義される。

(33)

特表平10-508391

$$\bar{x} = (\bar{x}_1^T, \bar{x}_2^T, \dots, \bar{x}_N^T)^T$$

$$\|\bar{x}\|^2 = \sum_{i=1}^N \|\bar{x}_i\|^2$$

これは、ラプラス密度を次のように与える。

ラブラシアン：

$$\|\bar{x} - \bar{y}\|_1 = \|\bar{x}\|_1 + \|\bar{y}\|_1 - \sum_{i=1}^N (2(1 - \frac{h(\bar{x}_i, \bar{y}_i)}{D}) \min(\|\bar{x}_i\|_1, \|\bar{y}_i\|_1))$$

この式において、 $h(\bar{x}_i, \bar{y}_i)$ が部分ベクトル \bar{x}_i 及び \bar{y}_i のハミング距

離である。前述したように、ベクトルユニットについてのテーブルを用いてハミング距離を計算する場合は、そのテーブルについて、部分ベクトルについて用いられるのと同じサイズのベクトルユニットを用いることが望ましい。

他の実施例においては、観察ベクトルについて基準ベクトルの部分集合が選択された後、 $N \geq 1$ として N 個の連続する観察ベクトルについて同一の部分集合が用いられる。従って、同一の部分集合が $N + 1$ 個の観察について用いられる。 $N + 1$ 個の観察に続く観察については、基準ベクトルの新しい部分集合が選択される。このメカニズムは続く観察で繰り返される。数 N は実験によって選択することができる。大語彙連続音声認識のためのフィリップス研究システムでは、 $N = 1$ (2つの連続する観察ベクトルが同一の部分集合を用いる) を用いて良好な結果を得ている。これに代えて、以前の観察ベクトルについて決定された部分集合を維持するのか、又は、現在の観察ベクトルについて新たな部分集合を決定するのかを、図3の局所化装置50が、予め定められた数 N を用いて動的に決定することが好ましい。局所化装置50は、観察ベクトル \bar{o} の動的な挙動に基づいて決定を行うことが望ましい。例えば、局所化装置50は、部分集合が既に計算された観察ベクトル \bar{o} を記憶することができる。後続の観察ベクトルが記憶された観察ベク

トルに近い (即ち、距離が予め定められたしきい値以下である) 場合に限り、同一の部分集合を維持する。後続の観察ベクトルが記憶された観察ベクトルから遠

(34)

特表平10-508391

く離れている場合は、新しい部分集合を決定する。

他の実施例においては、局所化装置が、各基準ユニットについてその基準ユニットを代表する少なくとも1つの基準ベクトルが基準ベクトルの部分集合の一員であるようにする。このため、局所化装置50は、観察ベクトルについて最初の部分集合を選択した後、各基準ユニットが最初の部分集合中で表現されているか否かを検証する。表現されていない場合は、局所化装置50が、表現されなかった基準ユニットについて、その基準ユニットを表現する基準ベクトルを加える。各基準ユニットについて、学習中に代表基準ベクトルを選択することができる。例え

ば、基準ベクトルは、基準ユニットを表現する全ての基準ベクトルの重み付け平均に最も近いものが選択される。代表基準ベクトルが前もって選択される場合には、このベクトルが基準パターンデータベース40に格納されることが望ましい。

他の実施例においては、代表基準ベクトルが動的に選択される。観察ベクトルについての最初の部分集合中で基準ユニットが表現されていない場合、以前の観察ベクトルについて基準ユニットを最良に表現することが見出された基準ベクトル、即ち、以前の観察ベクトルに対して最も小さい距離を持つ基準ベクトルが、代表基準ベクトルとして追加される。このため、観察ベクトルについて最初の部分集合を選択した後、局所化装置50が、この部分集合中で各基準ユニットが表現されるか否かをチェックする。表現されない場合は、表現されない基準ユニットについての代表基準ベクトルが追加される。次に、観察ベクトルと部分集合の各基準ベクトルとの全距離を計算し、次いで各基準ユニットについて、基準ユニットを表現し且つ観察ベクトルに最小の距離を有する基準ベクトルを選択する。この基準ベクトルが代表基準ベクトルとして選択され、例えば次の観察ベクトルのための部分集合が必要な場合に利用される。多数の連続する観察ベクトルについて代表基準ベクトルが同一であることに注意すべきである。各基準ユニットについて動的な方法を用いると、初期代表基準ベクトルを選択する必要があり、更に必要ならば最初の観察ベクトルを実現するために利用する。原理的には、基準ユニットを表現するいずれの基準ベクトルでも、初期代表基準ベクトルとして選択することができる。しかし、初期代表基準ベクトルとして、基準ユニットを表現

(35)

特表平10-508391

する全ての基準ベクトルの重み付け平均に最も近い基準ベクトルを選択することが好ましい。

次に、パターン認識におけるキーステップであり且つ本発明の主題である、1つの観察ベクトルと基準ベクトルとの距離の決定について説明する。連続的な物理量から導出される時系列パターンを認識するため、このキー要素を例えば遮蔽マルコフモデルのような他の技術と組合せて用いる方法については既によく知られている。各観察ベクトルについてこの技術を用いると、観察ベクトルと部分集合の各基準ベクトルとの間の、例えば尤度のようなベクトル類似度を計算することができる。各基準パターンについて、基準パターンを表現する基準ベクトルの

ベクトル類似度を組合せてパターン類似度を形成する。これを、以後の観察ベクトルに対して繰り返す。例えば最大尤度であるようにパターン類似度が最適に計算された基準パターンが、認識されたパターンとして位置付けられる。この認識されたパターンの出力のために出力手段70が用いられる。これは、例えば逐語的フォーマットでスクリーン上に認識したパターンを表示し、認識したパターンをメモリーに格納し、又は認識したパターンを例えば次の処理操作のためのコマンドのような入力として用いるような、種々の形をとることができる。例えばレベル化する方法のような技術を、基準ベクトルより多い観察ベクトルのシーケンスを含むパターン認識に利用できることが既によく知られている。例えば、ワード又はセンテンス全体を認識するためにサブワードユニットを基準パターンとして用いることが既知である。パターン認識に例えば発音辞書及び文法のような付加的な制約を設けてもよいことも、よく理解されている。発音辞書のような付加的な情報を、基準パターンデータベースを格納するために用いるものと同じのメモリーを用いて格納することができる。

上述では主として音声認識について説明したが、本発明は、時系列的性質を持ち且つ連続的な物理量から導出されるパターンのパターン認識ならばどのような型にも適用することができる。

本発明は、数学的用語で以下のように要約することができる。有限の数である

(36)

特表平10-508391

K個の基準ベクトル $\bar{\mu}_1$ 乃至 $\bar{\mu}_K$ が与えられ、全体として基準ベクトルの集合 P
 $= \{\bar{\mu}_1, \dots, \bar{\mu}_K\}$ を形成する。これらのベクトルの成分は通常連続値(R の要素

)を表し、 $P \subset R^D$ になる。ここで D は基準ベクトルのディメンションである。
 実際には、連続値ではなく、典型的には2バイト整数が用いられることに注意す
 べきである。タスクは、効率的な方法で、各入力観察ベクトル $\bar{o} \in O$ (O は可能
 な観察ベクトルの集合であり、 $O = R^D$ である)について、観察ベクトル \bar{o} に最
 も近い基準ベクトル $\bar{\mu}^{opt}$ を見付けることである。即ち、

$$\bar{\mu}^{opt} = \underset{\bar{\mu} \in P}{\operatorname{argmin}} d(\bar{\mu}, \bar{o})$$

である。集合 P を高い尤度で $\bar{\mu}^{opt}$ を含む P の部分集合に減らし、且つそのより
 小さい部分集合で $\bar{\mu}^{opt}$ をサーチすることにより、計算上複雑な距離の計算を減
 らすことができる。部分集合は各観察ベクトル \bar{o} について選択される。部分集合
 関数 s は、各観察ベクトル \bar{o} について P の部分集合 P_s を与える関数として定義

することができる。即ち、 $s: O \rightarrow G(P)$ である。ここで、 $G(P)$ は、集合
 P のパワー集合(全ての部分集合の集合)である。これにより、少ないタスクで
 基準ベクトル $\bar{\mu}^*$ を発見することが可能になる。即ち

$$\begin{aligned} \bar{\mu}^* &= \underset{\bar{\mu} \in P_s}{\operatorname{argmin}} d(\bar{\mu}, \bar{o}) \\ P_s &= s(\bar{o}) \\ P_s &\subset P \end{aligned}$$

である。部分集合 P_s が $\bar{\mu}^{opt}$ を含まないことがあり得る。関数 s は、ほぼ $\bar{\mu}^*$
 $= \bar{\mu}^{opt}$ になるようにし、 $\bar{\mu}^{opt}$ を含む $G(P)$ の尤度が大きくなるように設計

される。

E. Bocchieri "Vector quantization for the efficient computation of continuous likelihoods", Proceeding of ICASSP, 1993年、第 692-695頁から既知の方法
 においては、最初基準ベクトルを近隣にクラス化する。各近隣はコードワー

(37)

特表平10-508391

\overline{v}_i によって表示される。 $i = 1 \dots N$ である。コードワード \overline{v}_i の集合 V は実質的に集合 P より小さい。コードワード \overline{v}_i は P から導出される。各コードワード \overline{v}_i について、対応する P の部分集合 P_i は前もって決定される。認識の間、通常のコードワード量子化を用いて、観測ベクトル \overline{o} がコードワード $\overline{v}_i \in V$ に量子化される。 \overline{v}_i に対応して予め定められた部分集合 P_i は、 \overline{o} についての部分集合として用いられる。

クラスタ化： $P \rightarrow V = \{\overline{v}_i\}$ 及び $\overline{v}_i \rightarrow P_i \in G(P)$

コードワード量子化： $\overline{o} \rightarrow \overline{v}_i$

結果： $s : O \rightarrow V \rightarrow G(P)$

本発明によれば、基準ベクトル $\overline{\mu}$ 及び観測ベクトル \overline{o} を量子化するために量子化関数 \overline{R} が用いられる。従って、 \overline{R} はマッピング $P \rightarrow P_R$ 及び $O \rightarrow O_R$ を帰納する。量子化観測ベクトル $\overline{R}(\overline{o})$ と各量子化基準ベクトル $\overline{R}(\overline{\mu})$ との間の距離 $d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu}))$ を計算する。これらの距離に基づいて、量子化基準ベクトルの部分集合を選択する。これは、 P_R 空間中で部分集合関数 s_R によって表すことができる。即ち、

$s_R : O_R \rightarrow G(P_R)$

$s_R(\overline{R}(\overline{o})) =$

$\{\overline{R}(\overline{\mu}) \in P_R : d(\overline{R}(\overline{o}), \overline{R}(\overline{\mu})) < t(\overline{o}, P)\}$

である。しきい値関数 $t(\overline{o}, P)$ はしきい値を与える。これに代えて、選択を例えば予め定められた数又は比率のような他の機構に基づいて行ってもよい。例えば量子化が更に逆マップ $\overline{R}^* : G(P_R) \rightarrow G(P)$ を帰納するように、量子

化基準ベクトルの各部分集合について基準ベクトルの対応する部分集合を決定することができる。即ち、

$\overline{R}^*(M) = \{\overline{\mu} \in P : \overline{R}(\overline{\mu}) \in M\}$

である。実際には、各基準ベクトルとその量子化基準ベクトルとの間に、典型的には固定された関係が保持される。これは、次の図に示すように s を定める。

(38)

特表平10-508391

$$\begin{array}{ccc}
 \mathbf{g} : \mathbf{O} & \xrightarrow{\mathbf{S}} & \mathbf{G}(\mathbf{P}) \\
 \downarrow \overline{\mathbf{R}} & & \downarrow \overline{\mathbf{R}}^* \\
 \mathbf{O}_{\overline{\mathbf{R}}} & \xrightarrow{\mathbf{S}^*} & \mathbf{G}(\mathbf{P}_{\overline{\mathbf{R}}})
 \end{array}$$

式において、 \mathbf{s} は次のように与えられる。

$$\mathbf{s}(\overline{\mathbf{o}}) = \{\overline{\mathbf{u}} \in \mathbf{P} : d(\overline{\mathbf{R}}(\overline{\mathbf{o}}), \overline{\mathbf{R}}(\overline{\mathbf{u}})) < t(\overline{\mathbf{o}}, \mathbf{P})\}$$

基準ベクトルの量子化は前もって（例えば、学習の一部として）行うことができ、且つ観察ベクトルの量子化は1度行う必要があるだけであるので、量子化ベクトル間の距離の計算が元のベクトル間の計算に比べて簡単である場合は、一般的に計算上の節約が達成される。

これを達成するための1つの方法は、ベクトル成分をより小さい個別の集合 \mathbf{B} に量子化して $\overline{\mathbf{R}} : \mathbf{R}^D \rightarrow \mathbf{B}^D$ を与えることにより、ベクトルの成分の複雑さを減らす量子化関数を用いることである。或る種のプロセッサにおいては、例えばこの節約は、ベクトル成分を2バイト値（連続値を表す）から1バイト値に減らすことによって行うことができる。既に見たように、例えば $\mathbf{B} = \{-1, 1\}$ （これは $\{0, 1\}$ と表すこともできる）のようにベクトル成分を二値の値に減らすことによって、かなりの節約を達成することができる。このような量子化関数の一例

は $\overline{\mathbf{R}}(\overline{\mathbf{x}}) = (\text{符号}(x_1), \dots, \text{符号}(x_D))^T$ である。量子化ベクトルとして二値ベクトルと例えば元のベクトルのノルムのようなスカラー量との積を用いることにより、良い結果が得られている。これは、量子化関数 $\overline{\mathbf{R}}(\overline{\mathbf{x}}) = (\text{符号}(x_1), \dots, \text{符号}(x_D), \|\overline{\mathbf{x}}\|)^T$ が、 $D+1$ 次元ベクトルについての

特別な距離尺度を用いて、マッピング $\mathbf{R} : \mathbf{R}^D \rightarrow \mathbf{B}^D \times \mathbf{R}^1$ を帰納すると見ることができる。前述したように、部分ノルムの利用により、精度を更に向上させることができる。この一例は、 $D+F$ 次元ベクトルについての特別な距離尺度を用いて、マッピング $\mathbf{R} : \mathbf{R}^D \rightarrow \mathbf{B}^D \times \mathbf{R}^F$ を帰納する量子化関数 $\overline{\mathbf{R}}(\overline{\mathbf{x}}) = (\text{符号}(x_1), \dots, \text{符号}(x_D), \|\overline{\mathbf{x}}_1\|, \|\overline{\mathbf{x}}_2\|, \dots, \|\overline{\mathbf{x}}_F\|)^T$ を用いるこ

(39)

特表平10-508391

とである。

量子化の他の方法は、ベクトル成分の複雑さを同一に保ちながらベクトルのディメンションを減らし、 $\bar{R}: R^D \rightarrow R^F$ 又は $\bar{R}: R^D \rightarrow B^F$ とすることである。

ここで、 $F < D$ である。この例は、部分ノルムを用いるか又は符号ベクトル成分の部分集合を用いる。

両方法を1つに組合せて $\bar{R}: R^D \rightarrow B^F \times R^F$ とすることもできる。ここで、 $E \leq D$ 、 $F \leq D$ である。

【図1】

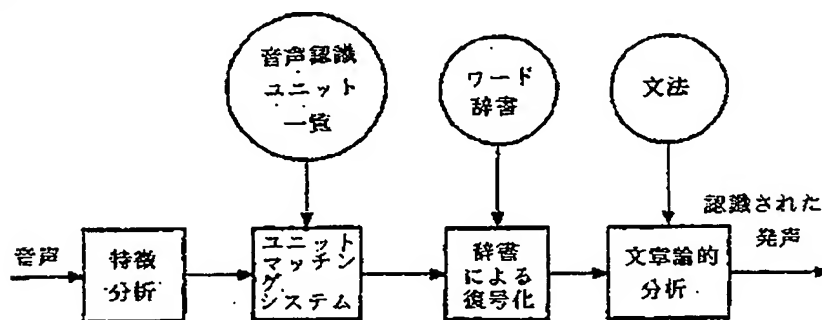


FIG. 1

【図2】

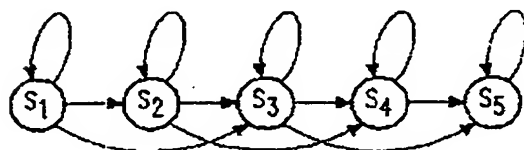


FIG. 2

(40)

特表平10-508391

【図3】

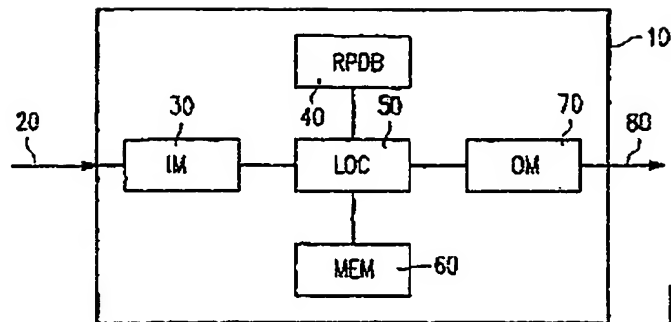


FIG. 3

【図6】

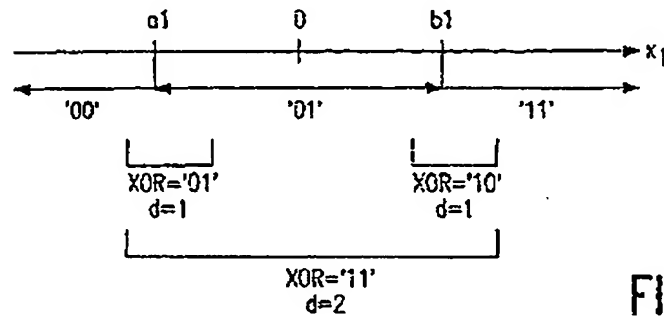


FIG. 6

【図4】

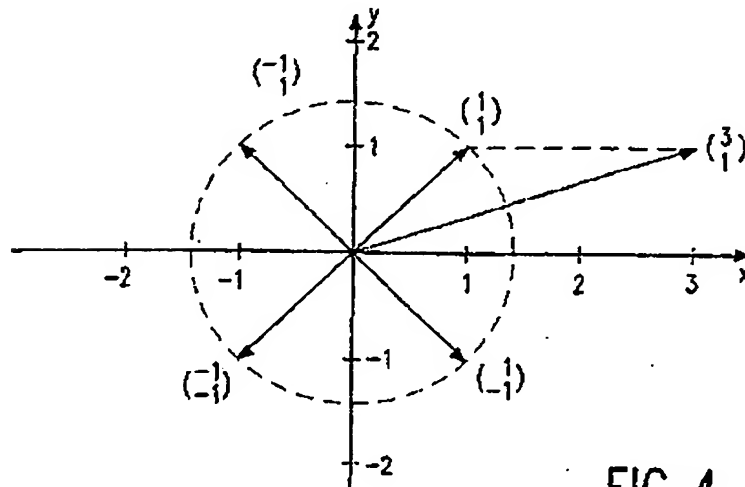


FIG. 4

(41)

特表平10-508391

【図7】

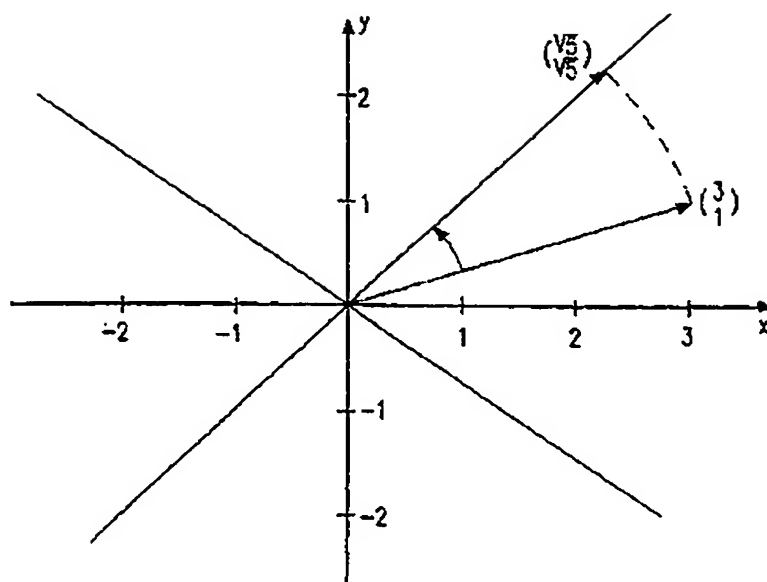


FIG. 7

(42)

特表平10-508391

【図5】

$$\bar{x} = \begin{pmatrix} 51 \\ -72 \\ 46 \end{pmatrix} ; \bar{S}(\bar{x}) = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

FIG. 5

i	\bar{y}_i	$\bar{x} - \bar{y}_i$	$\ \bar{x} - \bar{y}_i\ _1$	$\bar{S}(\bar{y}_i)$	$\bar{S}(\bar{x}) - \bar{S}(\bar{y}_i)$	$\ \bar{S}(\bar{x}) - \bar{S}(\bar{y}_i)\ _1$
0	$\begin{pmatrix} 23 \\ -31 \\ -90 \end{pmatrix}$	$\begin{pmatrix} 28 \\ -41 \\ 136 \end{pmatrix}$	205	$\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}$	2
1	$\begin{pmatrix} 66 \\ 21 \\ 28 \end{pmatrix}$	$\begin{pmatrix} -15 \\ -93 \\ 18 \end{pmatrix}$	126	$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2 \\ 0 \end{pmatrix}$	2
2	$\begin{pmatrix} 35 \\ 82 \\ -41 \end{pmatrix}$	$\begin{pmatrix} 16 \\ -154 \\ 87 \end{pmatrix}$	257	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}$	4
3	$\begin{pmatrix} -96 \\ 40 \\ 70 \end{pmatrix}$	$\begin{pmatrix} 147 \\ -112 \\ -24 \end{pmatrix}$	283	$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$	4
4	$\begin{pmatrix} 40 \\ 33 \\ -34 \end{pmatrix}$	$\begin{pmatrix} 11 \\ -105 \\ 80 \end{pmatrix}$	195	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}$	4
5	$\begin{pmatrix} -35 \\ 87 \\ 65 \end{pmatrix}$	$\begin{pmatrix} 86 \\ -159 \\ -19 \end{pmatrix}$	264	$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ -2 \\ 0 \end{pmatrix}$	4
6	$\begin{pmatrix} 19 \\ 60 \\ -13 \end{pmatrix}$	$\begin{pmatrix} 32 \\ -132 \\ 59 \end{pmatrix}$	223	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}$	4
7	$\begin{pmatrix} 61 \\ -5 \\ 8 \end{pmatrix}$	$\begin{pmatrix} -10 \\ -67 \\ 38 \end{pmatrix}$	115	$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	0
8	$\begin{pmatrix} 74 \\ -53 \\ 5 \end{pmatrix}$	$\begin{pmatrix} -23 \\ -19 \\ 41 \end{pmatrix}$	83	$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$	0
9	$\begin{pmatrix} -59 \\ -39 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 110 \\ -33 \\ 45 \end{pmatrix}$	188	$\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$	$\begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$	2

(43)

特表平10-508391

【図8】

$$\bar{x} = \begin{pmatrix} 51 \\ -72 \\ 46 \end{pmatrix} ; \bar{S}(\bar{x}) = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} ; \|\bar{x}\|_1 = 169 ; \bar{S}(\bar{x}) \cdot \frac{\|\bar{x}\|_1}{3} = \begin{pmatrix} 56\frac{1}{3} \\ -56\frac{1}{3} \\ 56\frac{1}{3} \end{pmatrix}$$

i	\bar{y}_i	$\bar{S}(\bar{y}_i)$	$\ \bar{y}_i\ _1$	$\bar{S}(\bar{y}_i) \cdot \frac{\ \bar{y}_i\ _1}{4}$	$\bar{S}(\bar{x}) \cdot \frac{\ \bar{x}\ _1}{4} - \bar{S}(\bar{y}_i) \cdot \frac{\ \bar{y}_i\ _1}{4}$	d_{xy}	$\ \bar{x} - \bar{y}_i\ _1$
0	$\begin{pmatrix} 23 \\ -31 \\ -90 \end{pmatrix}$	$\begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}$	144	$\begin{pmatrix} 48 \\ -48 \\ -48 \end{pmatrix}$	$\begin{pmatrix} 8\frac{1}{3} \\ -8\frac{1}{3} \\ 104\frac{1}{3} \end{pmatrix}$	121	205
1	$\begin{pmatrix} 66 \\ 21 \\ 28 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$	115	$\begin{pmatrix} 38\frac{1}{3} \\ 38\frac{1}{3} \\ 38\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} 18 \\ -96\frac{2}{3} \\ 18 \end{pmatrix}$	$132\frac{2}{3}$	126
2	$\begin{pmatrix} 35 \\ 82 \\ -41 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	158	$\begin{pmatrix} 52\frac{2}{3} \\ 52\frac{2}{3} \\ -52\frac{2}{3} \end{pmatrix}$	$\begin{pmatrix} 3\frac{2}{3} \\ -109 \\ 109 \end{pmatrix}$	$221\frac{2}{3}$	257
3	$\begin{pmatrix} -96 \\ 40 \\ 70 \end{pmatrix}$	$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$	206	$\begin{pmatrix} -68\frac{2}{3} \\ 68\frac{2}{3} \\ 68\frac{2}{3} \end{pmatrix}$	$\begin{pmatrix} 125 \\ -125 \\ -12\frac{1}{3} \end{pmatrix}$	$262\frac{1}{3}$	283
4	$\begin{pmatrix} 40 \\ 33 \\ -34 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	107	$\begin{pmatrix} 35\frac{2}{3} \\ 35\frac{2}{3} \\ -35\frac{2}{3} \end{pmatrix}$	$\begin{pmatrix} 20\frac{2}{3} \\ -92 \\ 92 \end{pmatrix}$	$204\frac{2}{3}$	195
5	$\begin{pmatrix} -35 \\ 87 \\ 65 \end{pmatrix}$	$\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$	187	$\begin{pmatrix} -62\frac{1}{3} \\ 62\frac{1}{3} \\ 62\frac{1}{3} \end{pmatrix}$	$\begin{pmatrix} 118\frac{2}{3} \\ -118\frac{2}{3} \\ -6 \end{pmatrix}$	$243\frac{1}{3}$	264
6	$\begin{pmatrix} 19 \\ 60 \\ -13 \end{pmatrix}$	$\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$	92	$\begin{pmatrix} 30\frac{2}{3} \\ 30\frac{2}{3} \\ -30\frac{2}{3} \end{pmatrix}$	$\begin{pmatrix} 25\frac{2}{3} \\ -87 \\ 87 \end{pmatrix}$	$199\frac{2}{3}$	223
7	$\begin{pmatrix} 61 \\ -5 \\ 8 \end{pmatrix}$	$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$	74	$\begin{pmatrix} 24\frac{2}{3} \\ -24\frac{2}{3} \\ 24\frac{2}{3} \end{pmatrix}$	$\begin{pmatrix} 31\frac{2}{3} \\ -31\frac{2}{3} \\ 31\frac{2}{3} \end{pmatrix}$	95	115
8	$\begin{pmatrix} 74 \\ -53 \\ 5 \end{pmatrix}$	$\begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$	132	$\begin{pmatrix} 44 \\ -44 \\ 44 \end{pmatrix}$	$\begin{pmatrix} 12\frac{1}{3} \\ -12\frac{1}{3} \\ 12\frac{1}{3} \end{pmatrix}$	37	83
9	$\begin{pmatrix} -59 \\ -39 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}$	99	$\begin{pmatrix} -33 \\ -33 \\ 33 \end{pmatrix}$	$\begin{pmatrix} 89\frac{1}{3} \\ -23\frac{1}{3} \\ 23\frac{1}{3} \end{pmatrix}$	136	188

FIG. 8

(44)

特表平10-508391

【国際調査報告】

INTERNATIONAL SEARCH REPORT

International application No.
PCT/IS 96/00849

A. CLASSIFICATION OF SUBJECT MATTER		
IPC6: G10L 5/06, G06K 9/00 According to International Patent Classification (IPC) or to both national classification and IPC		
B. FIELDS SEARCHED		
Minimum documentation searched (classification system followed by classification symbols)		
IPC6: G10L		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
SE,OK,FI,NO classes as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
EPAT, INSPEC, JAPIC		
C. DOCUMENTS CONSIDERED TO BE RELEVANT		
Category ^a	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	EP 0615227 A2 (INTERNATIONAL BUSINESS MACHINES CORPORATION), 14 Sept 1994 (14.09.94), page 3, line 51 - page 5, line 28, claim 5	1-2, 10-14, 21-23
A	—	3-9, 15-20
A	IEEE Trans. on acoustics, speech and signal processing, Volume 29, No 2, April 1981, S.Furai, "Cepstral analysis technique for automatic speaker verification", page 255 - page 269, see pp. 255-259	1-23
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
^a Special categories of cited documents: ^{"A"} documents defining the general state of the art which is not considered to be of particular relevance ^{"B"} prior documents but published on or after the international filing date ^{"C"} documents which may have priority claim(s) or which is cited to establish the publication date of another document or other special reasons (as specified) ^{"O"} documents referring to an oral disclosure, use, exhibition or other means ^{"P"} documents published prior to the international filing date but later than the priority date claimed ^{"T"} later documents published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention ^{"X"} documents of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone ^{"Y"} documents of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is considered with one or more other such documents, such combinations being obvious to a person skilled in the art ^{"Z"} document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
26 February 1997		27-02-1997
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86		Authorized officer Lars Christerson Telephone No. +46 8 782 23 00

Form PCT/ISA/210 (second sheet) (July 1992)

(45)

待表平10-508391

INTERNATIONAL SEARCH REPORT

International application No.

PCT/IB 95/00849

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	Proc. on seventh international conference on pattern recognition, Volume 2, 1984, G.Rigoll et al, "Some tools for speaker-independent, isolated word recognition systems based on system theory and system dynamics algorithms", pp.1248-1250 --	1-23
A	EP 0627726 A1 (NEC CORPORATION), 7 December 1994 (07.12.94), page 2, line 1 - page 5, line 19; page 8, line 30 - page 9, line 19, Cited in the application --	1-23
A	Patent Abstracts of Japan. Vol 8, No 51, P-259, abstract of JP.A.58-200379 (MITSUBISHI DENKI K.K.), 21 November 1983 (21.11.83) --	1-23
A	Int. conference on acoustics speech and signal processing, Volume 2, 1993, E. Boechieri, "Vector quantization for the efficient computation of continuous density likelihoods. pp.II-692--II-695", page 692 - page 693, Cited in the application -----	1-23

Form PCT/ISA/210 (continuation of second sheet) (July 1993)

(46)

特表平 10-508391

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/IB 96/00849

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP-A2- 0615227	14/09/94	JP-A- 6274200 US-A- 5497447	30/09/94 05/03/96
EP-A1- 0627726	07/12/94	JP-B- 2522154 JP-A- 6348292	07/08/96 22/12/94

Form PCT/IS-A/210 (patent family annex) (July 1992)

METHOD AND SYSTEM FOR PATTERN RECOGNITION BASED ON DYNAMICALLY CONSTRUCTING A SUBSET OF REFERENCE VECTORS

Publication number: WO9708685

Publication date: 1997-03-06

Inventor: BEYERLEIN PETER; ULLRICH MEINHARD DIETER

Applicant: PHILIPS ELECTRONICS NV (NL); PHILIPS NORDEN AB (SE); PHILIPS PATENTVERWALTUNG (DE)

Classification:

- **International:** G10L15/08; G10L15/14; G10L15/00; (IPC1-7): G10L5/06

- **European:** G10L15/08

Application number: WO1996IB00849 19960826

Priority number(s): EP19950202317 19950828

Also published as:

WO9708685 (A3)
EP0788648 (A3)
EP0788648 (A2)
US5933806 (A1)
EP0788648 (A0)

more >>

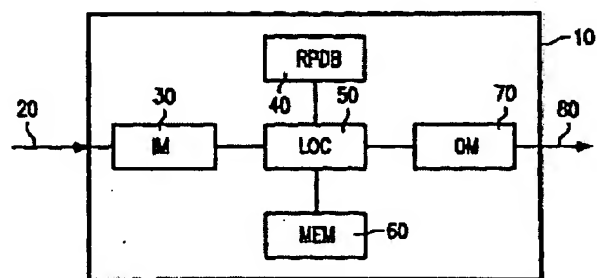
Cited documents:

EP0615227
EP0627726
JP58200379

Report a data error here

Abstract of WO9708685

The system and method are used for recognising a time-sequential input pattern (20), which is derived from a continual physical quantity, such as speech. The system comprises input means (30), which accesses the physical quantity and therefrom generates a plurality of input observation vectors. The input observation vectors represent the input pattern. A reference pattern database (40) is used for storing a plurality of reference patterns. Each reference pattern consists of a sequence of reference units, where each reference unit is represented by at least one associated reference vector $\langle o \rangle_{\mu a \langle o \rangle}$ in a set $\{\langle o \rangle_{\mu a \langle o \rangle}\}$ of reference vectors. A localizer (50) is used for locating among the reference patterns stored in the reference pattern database (40) a recognised reference pattern, which corresponds to the input pattern.



Data supplied from the esp@cenet database - Worldwide

METHOD AND SYSTEM FOR PATTERN RECOGNITION BASED ON DYNAMICAL CONSTRUCTING A SUBSET OF REFERENCE VECTORS

Publication number: EP0788648

Publication date: 1997-08-13

Inventor: BEYERLEIN PETER (DE); ULLRICH MEINHARD DIETER (DE)

Applicant: PHILIPS ELECTRONICS NV (NL); PHILIPS PATENTVERWALTUNG (DE)

Classification:

- **international:** G10L15/08; G10L15/14; G10L15/00; (IPC1-7): G10L5/06

- **europaen:** G10L15/08

Application number: EP19960926536 19960826

Priority number(s): EP19950202317 19950828; WO1996IB00849 19960826

Also published as:



WO9708685 (A3)
WO9708685 (A2)
EP0788648 (A3)
US5933806 (A1)
EP0788648 (A0)

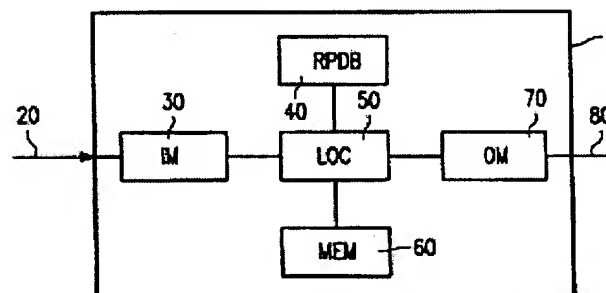
more >>

Report a data error he

Abstract not available for EP0788648

Abstract of corresponding document: **US5933806**

A system and method are used for recognising a time-sequential input pattern (20), which is derived from a continual physical quantity, such as speech. The system has input means (30), which accesses the physical quantity and therefrom generates a plurality of input observation vectors. The input observation vectors represent the input pattern. A reference pattern database (40) is used for storing a plurality of reference patterns. Each reference pattern includes a sequence of reference units, where each reference unit is represented by at least one associated reference vector +E, ovs mu +EE a in a set {+E, ovs mu +EE a} of reference vectors. A localizer (50) is used for locating among the reference patterns stored in the reference pattern database (40), a recognised reference pattern, which corresponds to the input pattern. The locating includes selecting a subset {+E, ovs mu +EE s} of reference vectors from said set {+E, ovs mu +EE a} for each input observation vector +E, ovs o+EE by calculating a measure of dissimilarity between the input observation vector +E, ovs o+EE and each reference vector of the set {+E, ovs mu +EE a}. A number of reference vectors +E, ovs mu +EE a, whose measures of dissimilarity with said input observation vector +E, ovs o+EE are the smallest, are used as the subset {+E, ovs mu +EE s} of reference vectors. The reference vectors of the subset {+E, ovs mu +EE s} are used to calculate for each reference pattern, a pattern similarity score. The recognised pattern is one of the reference patterns for which an optimum of the pattern similarity scores is calculated. Output means (70) are used for outputting the recognised pattern.



Data supplied from the esp@cenet database - Worldwide

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record.**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☐ BLACK BORDERS
- ☐ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☒ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☐ GRAY SCALE DOCUMENTS
- ☐ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☐ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.